

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:03:00 ; Search time 30.5 Seconds
(without alignments)
56.795 Million cell updates/sec

Title: US-09-787-986A-1
Perfect score: 83
Sequence: 1 NGVCCGYKCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_101002:*
1: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1980.DAT:*
2: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1981.DAT:*
3: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1982.DAT:*
4: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1983.DAT:*
5: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1984.DAT:*
6: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1985.DAT:*
7: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1986.DAT:*
8: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1987.DAT:*
9: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1988.DAT:*
10: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1989.DAT:*
11: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1990.DAT:*
12: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1991.DAT:*
13: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1992.DAT:*
14: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1993.DAT:*
15: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1994.DAT:*
16: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1995.DAT:*
17: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1996.DAT:*
18: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1997.DAT:*
19: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1998.DAT:*
20: /SIDB2/gcgdata/geneseq/genesep-emb1/AA1999.DAT:*
21: /SIDB2/gcgdata/geneseq/genesep-emb1/AA2000.DAT:*
22: /SIDB2/gcgdata/geneseq/genesep-emb1/AA2001.DAT:*
23: /SIDB2/gcgdata/geneseq/genesep-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	96.4	13	21	AA08017
2	80	96.4	13	21	AA092229
3	80	96.4	61	21	AA08016
4	80	96.4	61	21	AA092231
5	80	96.4	12	21	AA08018
6	74	89.2	13	21	AA092230
7	68	81.9	12	21	AA08019
8	66	79.5	13	21	AA08013
9	60	72.3	12	21	AA08014
10	60	72.3	12	21	AA08015

11	50	60.2	174	11	AA07664
12	46.5	56.0	28	23	ABB96853
13	46.5	56.0	28	23	ABB96857
14	46.5	56.0	74	23	ABB96598
15	46.5	56.0	74	23	ABB96598
16	46	55.4	55	22	ABB66559
17	46	55.4	69	21	AA083941
18	46	55.4	538	22	AA061111
19	46	55.4	971	22	AA097060
20	45.5	54.8	971	22	ABG24290
21	45	54.2	369	22	ABG15487
22	45	54.2	919	20	AA084970
23	44	53.0	437	20	AA032502
24	44	53.0	465	20	AA032501
25	43.5	52.4	109	22	AA063601
26	43	51.8	73	19	AA020881
27	43	51.8	136	23	ABP04912
28	43	51.8	164	21	AA061721
29	43	51.8	175	21	AA015202
30	43	51.8	175	23	AA097123
31	43	51.8	344	22	AA067521
32	43	51.8	451	22	AA067533
33	43	51.8	464	22	AA067517
34	43	51.8	464	22	AA067531
35	43	51.8	464	22	AA067532
36	43	51.8	485	23	AA048396
37	43	51.8	574	15	AA066034
38	43	51.8	574	15	AA066033
39	42	50.6	75	22	AA082164
40	42	50.6	130	22	AA039994
41	42	50.6	236	22	AA036921
42	42	50.6	246	22	AA058040
43	42	50.6	347	19	AA078103
44	42	50.6	598	23	ABP26696
45	42	50.6	623	22	ABB61371

ALIGNMENTS

RESULT 1	
AA08017	
ID	AA08017 standard; peptide: 13 AA.
XX	
AC	AA08017;
XX	
DT	14-NOV-2000 (first entry)
XX	
DE	Amino acid sequence of the conotoxin peptide Marl.
XX	
KW	Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.
XX	
OS	Conus marmoreus.
XX	
FT	Key
FT	Modified-site 12
FT	/note= "hydroxy-Pro"
XX	
PN	WO200044769-A1.
XX	
PD	03-AUG-2000.
XX	
PF	28-JAN-2000; 2000MO-US01978.
XX	
PR	29-JAN-1999; 9905-0118381.
XX	
PA	28-DEC-1999; 9905-017343.
XX	
PI	(UTAH) UNIV UTAH RES FOUNO.
XX	
DR	McIntosh JM, Olivera BM, Cruz LJ;
XX	WPI; 2000-476222/41.

Cysteine-added var
Omega-conopeptide
Omega-conopeptide
Omega-conopeptide
Omega-conopeptide
Drosophila melanog
Benzene ring hydro
Novel human polype
Human Ajbda-like p
Novel human diagno
Novel human diagno
Shrimp white spot
Human parkin gene
Human parkin gene
Human gastric can
Human presentin I
Human OREX protein
Human secreted pro
Granulopoietic act
Human G-CSF mutant
Amino acid sequenc
Amino acid sequenc
Amino acid sequenc
Amino acid sequenc
Amino acid sequenc
Human SSA-56kDa pr
Rat ARD 1. Rattus
Human ARD 1. Homo
Human haematologic
Propionibacterium
Human polypeptide,
Drosophila melanog
Chimeric receptor
Streptococcus poly
Drosophila melanog

PT Purified ap-conotoxin derived from cone snail venom for use as an
 XX analgesic -
 PS Claim 13; Page -: 29pp; English.
 CC The present sequence represents an ap-conotoxin peptide, designated
 CC Marl. Conotoxins are naturally available in minute amounts in the
 CC venom of cone snails. The peptides have analgesic activity. The
 CC peptides are used to treat or prevent pain.
 CC note: this sequence does not appear in the specification; it was created
 CC using information provided.
 XX
 SQ Sequence 13 AA;
 Query Match 96.4%; Score 80; DB 21; Length 13;
 Best Local Similarity 92.3%; Pred. No. 0.00029;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NGVCCGYKLCCHXC 13
 DB 1 NGVCCGYKLCCHPC 13
 RESULT 2
 AAY92229
 ID AAY92229 standard; peptide: 13 AA.
 XX
 AC AAY92229;
 XX
 DT 10-AUG-2000 (first entry)
 DE Chl-conotoxin peptide, chl-MrIA.
 XX
 DE Chl-conotoxin; chl-MrIA; cone snail; inhibitor; amine transporter;
 XX
 KM neuronal; noradrenergic transporter; urinary tract disorder; analgesic;
 KM antihypertensive; cardiac; antidepressant; anxiolytic; anti-inflammatory.
 XX
 OS Conus marmoreus.
 XX
 FT Key Location/Qualifiers
 FT Misc-difference 12 /Label= 4Hyp
 FT /note= "4-hydroxyproline"
 XX
 PN WO200020444-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 01-OCT-1999; 99WO-AU00844.
 XX
 PR 02-OCT-1998; 98AU-0006274.
 XX
 PA (UYOU) UNIV QUEENSLAND.
 XX
 PI Lewis RJ, Alewood PF, Sharpe JA;
 XX
 DR WPI: 2000-303738/26.
 XX
 PT Isolated, synthetic or recombinant chl-conotoxin peptide capable of
 PT inhibiting neuronal amine transporter used for treatment or prophylaxis
 PT of urinary or cardiovascular conditions, mood disorders, or
 PT treatment/control of pain/inflammation
 XX
 PS Claim 3; Page 33; 47pp; English.
 XX
 CC This conotoxin, chl-MrIA, is a member of a new class of conotoxins,
 CC designated chl-conotoxin. It was isolated from the venom of the mollusc
 CC hunting cone snail, Conus marmoreus. The peptide is an inhibitor of the
 CC neuronal amine transporters, especially the neuronal noradrenergic
 CC transporter. Inhibitors of noradrenergic re-uptake which have a
 CC negligible anti-cholinergic effect are particularly useful in the
 CC treatment of lower urinary tract disorders. Chl-MrIA (0.1 nM-1 micro M)
 CC inhibited the accumulation of radiolabeled noradrenergic in a

CC concentration-dependent manner, with a log IC-50 value of -8.17 plus or
 CC minus 0.0275 (n = 4). The concentration of chl-MrIA required to inhibit
 CC the accumulation by 50 percent was found to be approximately 7 nM. This
 CC concentration is approximately one order of magnitude lower than that
 CC needed for desipramine to produce the same effect. The peptides are
 CC useful for the treatment or prophylaxis of urinary or cardiovascular
 CC conditions or diseases (arrhythmia or coronary heart failure) or mood
 CC disorders (depression, anxiety or cravings), or the treatment or control
 CC of pain or inflammation (chronic pain, neuropathic pain or inflammatory
 CC pain).
 XX
 SQ Sequence 13 AA;
 Query Match 96.4%; Score 80; DB 21; Length 13;
 Best Local Similarity 92.3%; Pred. No. 0.00029;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NGVCCGYKLCCHXC 13
 DB 1 NGVCCGYKLCCHPC 13
 RESULT 3
 AAB08016
 ID AAB08016 standard; Protein: 61 AA.
 XX
 AC AAB08016;
 XX
 DT 14-NOV-2000 (first entry)
 DE Amino acid sequence of the conotoxin Marl propeptide.
 XX
 DE Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.
 XX
 KM Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.
 XX
 OS Conus marmoreus.
 XX
 PN WO200044769-A1.
 XX
 PD 03-AUG-2000.
 XX
 PF 28-JAN-2000; 2000WO-US01978.
 XX
 PR 29-JAN-1999; 99US-0118381.
 PR 28-DEC-1999; 99US-0173343.
 XX
 PA (UTAH) UNIV UTAH RES FOUND.
 XX
 PI McIntosh JM, Olivera BM, Cruz LJ;
 XX
 DR WPI: 2000-476222/41.
 DR N-PSDB; AAA63513.
 XX
 PT Purified ap-conotoxin derived from cone snail venom for use as an
 PT analgesic -
 XX
 PS Claim 23; Page 13-14; 29pp; English.
 XX
 CC The present sequence represents a Marl propeptide. Marl is an
 CC ap-conotoxin peptide. Conotoxins are naturally available in minute
 CC amounts in the venom of cone snails. The peptides have analgesic
 CC activity. The peptides are used to treat or prevent pain.
 XX
 SQ Sequence 61 AA;
 Query Match 96.4%; Score 80; DB 21; Length 61;
 Best Local Similarity 92.3%; Pred. No. 0.001; Mismatches 1; Indels 0; Gaps 0;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NGVCCGYKLCCHXC 13
 DB 49 NGVCCGYKLCCHPC 61

RESULT 4
 ID AAY92231 standard; Protein; 61 AA.
 AC AAY92231;
 DT 10-AUG-2000 (first entry)
 DE Chl-conotoxin, chl-MrIA, leader and mature peptide.
 DE Chl-conotoxin, chl-MrIA, cone snail; inhibitor; amine transporter;
 KW neuronal; noradrenergic transporter; urinary tract; analgesic; cardiant;
 KW antirhythmic; antidepressant; anxiolytic; anti-inflammatory.
 OS Conus marmoreus.
 PN WO200020444-A1.
 PD 13-APR-2000.
 PF 01-OCT-1999; 99WO-AU00844.
 PR 02-OCT-1998; 98AU-0006274.
 PA (UYOU) UNIV QUEENSLAND.
 PI Lewis RJ, Alewood PF, Sharpe IA;
 DR WPI, 2000-303738/26.
 DR N-PSDB; AAA09112.
 PT Isolated, synthetic or recombinant chl-conotoxin peptide capable of
 PT inhibiting neuronal amine transporter used for treatment or prophylaxis
 PT of urinary or cardiovascular conditions, mood disorders, or
 PT treatment/control of pain/inflammation
 PS Example 7; Page 31; 47pp; English.
 CC This sequence is the conotoxin, chl-MrIA, a member of a new class of
 CC conotoxins, designated chl-conotoxin. It was isolated from the venom of
 CC the mollusc hunting cone snail, Conus marmoreus. The peptide is an
 CC inhibitor of the neuronal amine transporters, especially the neuronal
 CC noradrenergic transporter. Inhibitors of noradrenergic re-uptake which
 CC have a negligible anti-cholinergic effect are particularly useful in the
 CC treatment of lower urinary tract disorders. Chl-MrIA (0.1 nM-1 micro M)
 CC inhibited the accumulation of radiolabeled noradrenaline in a
 CC concentration-dependent manner, with a log IC-50 value of -8.17 plus or
 CC minus 0.0275 (n = 4). The concentration of chl-MrIA required to inhibit
 CC the accumulation by 50 percent was found to be approximately 7 nM. This
 CC concentration is approximately one order of magnitude lower than that
 CC needed for desipramine to produce the same effect. The peptides are
 CC useful for the treatment or prophylaxis of urinary or cardiovascular
 CC conditions or diseases (arrhythmia or coronary heart failure) or mood
 CC disorders (depression, anxiety or cravings), or the treatment or control
 CC of pain or inflammation (chronic pain, neuropathic pain or inflammatory
 CC pain).
 CC
 SQ Sequence 61 AA:
 Query Match 96.4%; Score 80; DB 21; Length 61;
 Best Local Similarity 92.3%; Pred. No. 0.001;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGVCGYKLCCHXC 13
 DB 49 NGVCGYKLCCHPC 61

RESULT 5
 ID AAB08018 standard; peptide; 12 AA.
 AC AAB08018;
 PN

XX 14-NOV-2000 (first entry)
 DT Amino acid sequence of the conotoxin peptide Mar2.
 DE Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Mar2.
 KW Chl-conotoxin, chl-MrIB, leader and mature peptide.
 KW Chl-conotoxin, chl-MrIB, cone snail; inhibitor; amine transporter;
 KW neuronal; noradrenergic transporter; urinary tract; analgesic;
 KW antirhythmic; antidepressant; anxiolytic; anti-inflammatory.
 OS Conus marmoreus.
 PN WO200044769-A1.
 PD 03-AUG-2000.
 PF 28-JAN-2000; 2000WO-US01978.
 PR 29-JAN-1999; 99US-0118381.
 PR 28-DEC-1999; 99US-0173343.
 PA (UTAH) UNIV UTAH RES FOUND.
 PI McIntosh JM, Olivera BM, Cruz LJ;
 DR WPI, 2000-476222/41.
 PT Purified ap-conotoxin derived from cone snail venom for use as an
 PT analgesic -
 PS Claim 14; Page -; 29pp; English.
 CC The present sequence represents an ap-conotoxin peptide, designated
 CC Mar2. Conotoxins are naturally available in minute amounts in the
 CC venom of cone snails. The peptides have analgesic activity. The
 CC peptides are used to treat or prevent pain.
 CC note: this sequence does not appear in the specification; it was created
 CC using information provided.
 CC
 SQ Sequence 12 AA:
 Query Match 89.2%; Score 74; DB 21; Length 12;
 Best Local Similarity 91.7%; Pred. No. 0.0016;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVCYGYKLCCHXC 13
 DB 1 GVCYGYKLCCHPC 12

RESULT 6
 ID AAY92230 standard; peptide; 13 AA.
 AC AAY92230;
 DT 10-AUG-2000 (first entry)

DE Chl-conotoxin peptide, chl-MrIB.
 DE Chl-conotoxin, chl-MrIB, cone snail; inhibitor; amine transporter;
 KW neuronal; noradrenergic transporter; urinary tract disorder; analgesic;
 KW antirhythmic; antidepressant; anxiolytic; anti-inflammatory.
 OS Conus marmoreus.

Key Location/Qualifiers
 FT Misc-difference 12 /Label= 4Hyp
 FT /note= "4-hydroxyproline"
 FT
 PN WO200020444-A1.

XX 13-APR-2000.
PD 01-OCT-1999; 99WO-AU00844.
XX 02-OCT-1998; 98AU-0006274.
XX (UYOU) UNIV QUEENSLAND.
XX Lewis RJ, Alewood PF, Sharpe IA;
XX WPI; 2000-303738/26.
DR Isolated, synthetic or recombinant chl-conotoxin peptide capable of
XX inhibiting neuronal amine transporter used for treatment or prophylaxis
XX of urinary or cardiovascular conditions, mood disorders, or
XX treatment/control of pain/inflammation
XX
XX Claim 3; Page 33; 47pp; English.
XX This conotoxin, chl-Mr1B, is a member of a new class of conotoxins,
XX designated chl-conotoxin. It was isolated from the venom of the mollusc
XX hunting cone snail, Conus marmoreus. The peptide is an inhibitor of the
XX neuronal amine transporters, especially the neuronal noradrenaline
XX transporter. Inhibitors of noradrenaline re-uptake which have a
XX negligible anti-cholinergic effect are particularly useful in the
XX treatment of lower urinary tract disorders. Chl-Mr1A (0.1 nM-1 micro M)
XX inhibited the accumulation of radiolabeled noradrenaline in a
XX concentration-dependent manner, with a log IC-50 value of -8.17 plus or
XX minus 0.0275 (n = 4). The concentration of chl-Mr1A required to inhibit
XX the accumulation by 50 percent was found to be approximately 7 nM. This
XX concentration is approximately one order of magnitude lower than that
XX needed for desipramine to produce the same effect. The peptides are
XX useful for the treatment or prophylaxis of urinary or cardiovascular
XX conditions or diseases (arrhythmia or coronary heart failure) or mood
XX disorders (depression, anxiety or cravings), or the treatment or control
XX of pain or inflammation (chronic pain, neuropathic pain or inflammatory
XX pain).
XX
XX Sequence 13 AA;
SQ
Query Match 89.2%; Score 74; DB 21; Length 13;
Best Local Similarity 91.7%; Pred. No. 0.0017;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 GVCCGYKLCHXC 13
DB 2 GVCCGYKLCHPC 13
RESULT 7
AAB08019
ID AAB08019 standard; peptide; 12 AA.
XX AAB08019;
XX
XX 14-NOV-2000 (first entry)
XX Amino acid sequence of the conotoxin peptide U036.
XX Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; U036.
XX Conus marmoreus.
XX
XX Conus marmoreus.
XX
XX Key Location/Qualifiers
XX Misc-difference 7
XX /label= Lys, Xaa
XX /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
XX N,N,N-trimethyl Lys"
XX Modified-site 11
XX /note= "hydroxy-Pro"
XX
XX WO200044759-A1.

XX 03-AUG-2000.
PD 28-JAN-2000; 2000WO-US01978.
XX 29-JAN-1999; 99US-0118381.
XX 28-DEC-1999; 99US-0173343.
XX (UTAH) UNIV UTAH RES FOUND.
XX
XX McIntosh JM, Olivera BM, Cruz LJ;
XX WPI; 2000-476222/41.
DR Purified ap-conotoxin derived from cone snail venom for use as an
XX analgesic -
XX
XX Claim 15; Page -; 29pp; English.
XX The present sequence represents an ap-conotoxin peptide, designated
XX U036. Conotoxins are naturally available in minute amounts in the
XX venom of cone snails. The peptides have analgesic activity. The
XX peptides are used to treat or prevent pain
XX note: this sequence does not appear in the specification; it was created
XX using information provided.
XX
XX Sequence 12 AA;
SQ
Query Match 81.9%; Score 68; DB 21; Length 12;
Best Local Similarity 83.3%; Pred. No. 0.0099;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 GVCCGYKLCHXC 13
DB 1 GVCCGYKLCHPC 12
RESULT 8
AAB08013
ID AAB08013 standard; peptide; 13 AA.
XX AAB08013;
XX
XX 14-NOV-2000 (first entry)
XX Generic formula for conotoxin peptide Mar1.
XX
XX Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Mar1.
XX Conus marmoreus.
XX
XX Conus marmoreus.
XX
XX Key Location/Qualifiers
XX Misc-difference 7
XX /label= Tyr, Xaa
XX /note= "Xaa is mono-halo-Tyr, di-halo-Tyr,
XX O-sulfo-Tyr, O-phospho Tyr or nitro-Tyr"
XX
XX Misc-difference 8
XX /label= Lys, Xaa
XX /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
XX N,N,N-trimethyl Lys"
XX Misc-difference 12
XX /note= "optionally hydroxy-Pro"
XX
XX WO200044769-A1.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-US01978.
XX 29-JAN-1999; 99US-0118381.
XX 28-DEC-1999; 99US-0173343.
XX (UTAH) UNIV UTAH RES FOUND.
XX

XX McIntosh JM, Olivera BM, Cruz LJ;
XX venom of cone snails. The peptides have analgesic activity. The
DR WPI: 2000-476222/41.
XX Purified ap-conotoxin derived from cone snail venom for use as an
PT analgesic -
XX
PS Claim 2; Page 19; 29pp; English.
XX
CC The present sequence represents an ap-conotoxin peptide, designated
CC Mar1. Conotoxins are naturally available in minute amounts in the
CC venom of cone snails. The peptides have analgesic activity. The
CC peptides are used to treat or prevent pain.
XX
SQ Sequence 13 AA;
Query Match 79.5%; Score 66; DB 21; Length 13;
Best Local Similarity 76.9%; Pred. No. 0.019;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 NGVCGGYKLCHXC 13
DB 1 NGVCGGXKLCHPC 13
RESULT 9
AAB08014
ID AAB08014 standard; peptide: 12 AA.
XX
AC AAB08014;
XX
DF 14-NOV-2000 (first entry)
XX
DE Generic formula for conotoxin peptide Mar2.
XX
KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Mar2.
XX
OS Conus marmoreus.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /label= Tyr, Xaa
FT /note= "Xaa is mono-halo-Tyr, di-halo-Tyr,
FT O-sulfo-Tyr, O-phospho Tyr or nitro-Tyr"
FT Misc-difference 7 /label= Lys, Xaa
FT /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
FT N,N,N-trimethyl Lys"
FT Misc-difference 11 /note= "optionally hydroxy-Pro"
FT
XX WO200044769-A1.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-US01978.
XX
XX 29-JAN-1999; 99US-0118381.
XX
XX 28-DEC-1999; 99US-0173343.
XX
XX (UTAH) UNIV UTAH RES FOUND.
XX
XX McIntosh JM, Olivera BM, Cruz LJ;
XX
XX WPI: 2000-476222/41.
XX
XX Purified ap-conotoxin derived from cone snail venom for use as an
XX analgesic -
XX
PS Claim 2; Page 19; 29pp; English.
XX
CC The present sequence represents an ap-conotoxin peptide, designated

CC Mar2. Conotoxins are naturally available in minute amounts in the
CC venom of cone snails. The peptides have analgesic activity. The
CC peptides are used to treat or prevent pain.
XX
SQ Sequence 12 AA;
Query Match 72.3%; Score 60; DB 21; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GVCCGYKLCHXC 13
DB 1 GVCCGXKLCHPC 12
RESULT 10
AAB08015
ID AAB08015 standard; peptide: 12 AA.
XX
AC AAB08015;
XX
DF 14-NOV-2000 (first entry)
XX
DE Generic formula for conotoxin peptide U036.
XX
KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; U036.
XX
OS Conus marmoreus.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /label= Tyr, Xaa
FT /note= "Xaa is mono-halo-Tyr, di-halo-Tyr,
FT O-sulfo-Tyr, O-phospho Tyr or nitro-Tyr"
FT Misc-difference 7 /label= Lys, Xaa
FT /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
FT N,N,N-trimethyl Lys"
FT Misc-difference 11 /note= "optionally hydroxy-Pro"
FT
XX WO200044769-A1.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-US01978.
XX
XX 29-JAN-1999; 99US-0118381.
XX
XX 28-DEC-1999; 99US-0173343.
XX
XX (UTAH) UNIV UTAH RES FOUND.
XX
XX McIntosh JM, Olivera BM, Cruz LJ;
XX
XX WPI: 2000-476222/41.
XX
XX Purified ap-conotoxin derived from cone snail venom for use as an
XX analgesic -
XX
PS Claim 2; Page 19; 29pp; English.
XX
XX The present sequence represents an ap-conotoxin peptide, designated
XX U036. Conotoxins are naturally available in minute amounts in the
XX venom of cone snails. The peptides have analgesic activity. The
XX peptides are used to treat or prevent pain.
XX
SQ Sequence 12 AA;
Query Match 72.3%; Score 60; DB 21; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GVCCGYKLCHXC 13

Db 1 GVCCGXKLCHPC 12

RESULT 11

AA07664

AA07664 standard; protein; 174 AA.

AA07664;

13-FEB-1991 (first entry)

Cysteine-added variant of Granulocyte-Colony Stimulating Factor.

cysteine-added variant; cancer chemotherapy;

granulocyte-colony stimulating factor (G-CSF).

Homo sapiens.

MO9012874-A.

01-NOV-1990.

19-APR-1990; 90WO-US02144.

21-APR-1989; 89US-0341990.

(GENE-) GENETICS INST INC.

Shaw G, Veldman G, Wooters J;

WPI; 1990-348477/46.

Cysteine-added variants of interleukin-3, erythropoietin etc. -

contain at least 1 cysteine residue attached to polyalkylene

glycol gp. for improved pharmacokinetic properties

Dislosure; Page 15; 46pp; English.

Ala at position 37 of mature, human G-CSF is replaced by Cys. The

Cys residue is attached to a polyalkylene glycol moiety

which enhances the pharmacokinetic properties of the protein. In

addition, Cys at position 17 is replaced by Ala to prevent possible

improper disulphide bridge formation. The Thr and Pro residues at

positions 1 and 2 of native G-CSF (= positions 2 and 3 of this

sequence) may also be deleted to permit more consistent removal

of the N-terminal methionine.

See also AA07663, AA07665 and AA08358-9.

Sequence 174 AA:

Query Match 60.2%; Score 50; DB 11; Length 174;

Best Local Similarity 77.8%; Pred. No. 20;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

3 VCCGYKLCH 11

35 LCCTYKLCH 43

Omega-conopeptide J414 toxin sequence.

Omega-conopeptide; analgesic; anticonvulsant; vasotropic; cardiant;

neuroprotective; cerebroprotective; cardiovascular; antiinflammatory;

antimigraine; antidiabetic; tranquiliser; vulnerary; antipsychotic;

anxiolytic; neuroleptic; voltage gated ion channel; seizure; epilepsy;

neurological disorder; neurotoxic injury; hypoxia; anoxia; ischaemia;

stroke; cerebrovascular accident; brain trauma; spinal chord trauma;

drowning; suffocation; perinatal asphyxia; hypoglycaemic event; pain;

migraine; inflammation; cardiovascular disorder; psychiatric disorder;

psychosis; anxiety; schizophrenia.

Conus sp.

MO200207675-A2.

31-JAN-2002.

23-JUL-2001; 2001WO-US23041.

21-JUL-2000; 2000US-219616P.

05-FEB-2001; 2001US-265888P.

(UTAH) UNIV UTAH RES FOUND.

(COGN-) COGNETIX INC.

Olivera BM, McIntosh JM, Watkins M, Garrett JE, Shon K;

Jacobsen R, Jones RM, Cartier GE;

WPI; 2002-257318/30.

New omega-conopeptides useful for treating disorders associated with

voltage gated ion channels e.g. pain, inflammation, neurological or

cardiovascular disorders -

Claim 1(a); Page 72; 195pp; English.

The invention relates to isolated omega-conopeptides, nucleic acid

sequences encoding them, and propeptide sequences. The activity of

the peptides of the invention may be described as, analgesic,

anticonvulsant, vasotropic, cardiant, neuroprotective, cerebroprotective,

cardiovascular, antiinflammatory, antimigraine, antidiabetic,

tranquilliser, vulnerary, antipsychotic, anxiolytic and neuroleptic.

Peptides of the invention act by modulating the activity of voltage gated

ion channels. They may be used for treating or preventing disorders

associated with voltage gated ion channels such as neurological

disorders, e.g. seizure (associated with epilepsy), neurotoxic injury

associated with conditions of hypoxia, anoxia, ischaemia, stroke,

cerebrovascular accident, brain or spinal chord trauma, drowning,

suffocation, perinatal asphyxia or hypoglycaemic events; pain e.g.

migraine; inflammation or cardiovascular disorders. They may also be used

for treating psychiatric disorders e.g. psychosis, anxiety or

schizophrenia. The analgesic agents of the invention show diminished side

effects and toxicity, and are non-addictive. The sequences given in

records ABB96807-ABB96905 represent omega-conopeptide toxin sequences.

Sequence 28 AA:

Query Match 56.0%; Score 46.5; DB 23; Length 28;

Best Local Similarity 56.2%; Pred. No. 13;

Matches 9; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

1 NGVCCGY-----KLCH 11

10 NRVC CGYCSKRTKLCH 25

Omega-conopeptide I6.2 toxin sequence.

Omega-conopeptide; analgesic; anticonvulsant; vasotropic; cardiant;

XX	neuroprotective; cerebroprotective; cardiovascular; antiinflammatory;
KW	antimigraine; antidiabetic; tranquiliser; vulnerary; antipsychotic;
KW	anxiolytic; neuroleptic; voltage gated ion channel; seizure; epilepsy;
KW	neurological disorder; neurotoxic injury; hypoxia; anoxia; ischaemia;
KW	stroke; cerebrovascular accident; brain trauma; spinal chord trauma;
KW	drowning; suffocation; perinatal asphyxia; hypoglycaemic event; pain;
KW	migraine; inflammation; cardiovascular disorder; psychiatric disorder;
KW	psychosis; anxiety; schizophrenia.
XX	
OS	Conus lynceus.
XX	
PN	WO200207675-A2.
XX	
PD	31-JAN-2002.
XX	
PF	23-JUL-2001; 2001WO-US23041.
XX	
PR	21-JUL-2000; 2000US-219616P.
XX	
PR	05-FEB-2001; 2001US-265888P.
XX	
PA	(UTAH) UNIV UTAH RES FOUND.
XX	(COGN -) COGNETIX INC.
PI	Olivera BM, McIntosh JM, Watkins M, Garrett JE, Shon K;
XX	Jacobson R, Jones RM, Cartier GE;
DR	WPI; 2002-257318/30.
XX	
PT	New omega-conopeptides useful for treating disorders associated with
PT	voltage gated ion channels e.g. pain, inflammation, neurological or
PT	cardiovascular disorders -
XX	
PS	Claim 1(a): Page 72; 195pp; English.
XX	
CC	The invention relates to isolated omega-conopeptides, nucleic acid
CC	sequences encoding them, and propeptide sequences. The activity of
CC	the peptides of the invention may be described as, analgesic,
CC	anticonvulsant, vasotropic, cardiant, neuroprotective, cerebroprotective,
CC	cardiovascular, antiinflammatory, antimigraine, antidiabetic,
CC	tranquilliser, vulnerary, antipsychotic, anxiolytic and neuroleptic.
CC	Peptides of the invention act by modulating the activity of voltage gated
CC	ion channels. They may be used for treating or preventing disorders
CC	associated with voltage gated ion channels such as neurological
CC	disorders, e.g. seizure (associated with epilepsy), neurotoxic injury
CC	associated with conditions of hypoxia, anoxia, ischaemia, stroke,
CC	cerebrovascular accident, brain or spinal chord trauma, drowning,
CC	suffocation, perinatal asphyxia or hypoglycaemic events; pain e.g.
CC	migraine; inflammation or cardiovascular disorders. They may also be used
CC	for treating psychiatric disorders e.g. psychosis, anxiety or
CC	schizophrenia. The analgesic agents of the invention show diminished side
CC	effects and toxicity, and are non-addictive. The sequences given in
CC	records ABB96807-ABB96905 represent omega-conopeptide toxin sequences.
XX	
SO	Sequence 28 AA:
XX	
Query Match	56.0%; Score 46.5; DB 23; Length 28;
Best Local Similarity	56.2%; Pred. NO. 13;
Matches 9; Conservative	0; Mismatches 2; Indels 5; Gaps 1
OY	1 NGVCCGY-----KLCH 11
	1
	1
Db	10 NRVCCGYCKRPTLCH 25
XX	
RESULT 14	
ID	ABB96598 standard; Peptide; 74 AA.
XX	
AC	ABB96598;
XX	
DT	12-JUL-2002 (first entry)
XX	
DE	Omega-conopeptide J414 propeptide.

KW		Omega-conopeptide; analgesic; anticonvulsant; vasotropic; cardiant;
KW		neuroprotective; cerebroprotective; cardiovascular; antiinflammatory;
KW		antimigraine; antidiabetic; tranquiliser; vulnerary; antipsychotic;
KW		anxiolytic; neuroleptic; voltage gated ion channel; seizure; epilepsy;
KW		neurological disorder; neurotoxic injury; hypoxia; anoxia; ischaemia;
KW		stroke; cerebrovascular accident; brain trauma; spinal chord trauma;
KW		drowning; suffocation; perinatal asphyxia; hypoglycaemic event; pain;
KW		migraine; inflammation; cardiovascular disorder; psychiatric disorder;
KW		psychosis; anxiety; schizophrenia.
OS		Conus sp.
XX		
PN		WO200207675-A2.
XX		
PD		31-JAN-2002.
XX		
PE		23-JUL-2001; 2001WO-US23041.
XX		
PR		21-JUL-2000; 2000US-219616P.
PR		05-FEB-2001; 2001US-265888P.
XX		
PA		(UTAH) UNIV UTAH RES FOUND.
PA		(COGN-) COGNETIX INC.
XX		
PI		Olivera BM, McIntosh JM, Watkins M, Garrett JE, Shon K;
PI		Jacobsen R, Jones RM, Cartier GB;
XX		
DR		WPI: 2002-257318/30.
XX		N-PsDB; ABL96858.
PT		New omega-conopeptides useful for treating disorders associated with
PT		voltage gated ion channels e.g. pain, inflammation, neurological or
PT		cardiovascular disorders -
XX		
PS		Claim 1(c): Page 24; 195pp; English.
CC		The invention relates to isolated omega-conopeptides, nucleic acid
CC		sequences encoding them, and propeptide sequences. The activity of
CC		the peptides of the invention may be described as, analgesic,
CC		anticonvulsant, vasotropic, cardiant, neuroprotective, cerebroprotective,
CC		cardiovascular, antiinflammatory, antimigraine, antidiabetic,
CC		tranquilliser, vulnerary, antipsychotic, anxiolytic and neuroleptic.
CC		Peptides of the invention act by modulating the activity of voltage gated
CC		ion channels. They may be used for treating or preventing disorders
CC		associated with voltage gated ion channels such as neurological
CC		disorders, e.g. seizure (associated with epilepsy), neurotoxic injury
CC		associated with conditions of hypoxia, anoxia, ischaemia, stroke,
CC		cerebrovascular accident, brain or spinal chord trauma, drowning,
CC		suffocation, perinatal asphyxia or hypoglycaemic events; pain e.g.
CC		migraine; inflammation or cardiovascular disorders. They may also be used
CC		for treating psychiatric disorders e.g. psychosis, anxiety or
CC		schizophrenia. The analgesic agents of the invention show diminished side
CC		effects and toxicity, and are non-addictive. The sequences given in
CC		records ABB96595-ABB96697 represent omega-conopeptide propeptide
CC		sequences.
XX		
SQ		Sequence 74 AA:
Query Match	56.0%;	Score 46.5; DB 23; Length 74;
Best Local Similarity	56.2%;	Pred. No. 29;
Matches 9; Conservative	0;	Mismatches 2; Indels 5; Gaps 1.
OY	1	NGVCGGY-----KLCH 11
	I I I I I I I	I I I
DB	55	NRVCGGYSKRTTHLCH 70

Job time : 31.5 secs

XX 12-JUL-2002 (first entry)
DT
XX
DEXX Omega-conopeptide L6.2 propeptide.
DEXX Omega-conopeptide; analgesic; anticonvulsant; vasotropic; cardiant;
KW neuroprotective; cerebroprotective; cardiovascular; antiinflammatory;
KW antimigraine; antidiabetic; tranquiliser; vulnerary; antipsychotic;
KW anxiolytic; neuroleptic; voltage gated ion channel; seizure; epilepsy;
KW neurological disorder; neurotoxic injury; hypoxia; anoxia; ischaemia;
KW stroke; cerebrovascular accident; brain trauma; spinal chord trauma;
KW drowning; suffocation; perinatal asphyxia; hypoglycaemic event; pain;
KW migraine; inflammation; cardiovascular disorder; psychiatric disorder;
KW psychosis; anxiety; schizophrenia.
XXXX Conus leopardus.
OSXX W0200207675-A2.
PNXX 31-JAN-2002.
PDXX 23-JUL-2001; 2001WO-US23041.
PEXX 21-JUL-2000; 2000US-219616P.
PRXX 05-FEB-2001; 2001US-265888P.
PRXX (UTAH) UNIV UTAH RES FOUND.
PAXX (COGN-) COGNEXIX INC.
PAXX Oliveira BM, McIntosh JM, Watkins M, Garrett JE, Shon K;
PI Jacobsen R, Jones RM, Cartier GE;
PIXX MPI: 2002-257318/30.
DRXX N-PSDB: ABL98913.
DRXX New omega-conopeptides useful for treating disorders associated with
PT voltage gated ion channels e.g. pain, inflammation, neurological or
PT cardiovascular disorders -
PTXX Claim 1(c); Page 50; 195pp; English.
PSXX The invention relates to isolated omega-conopeptides, nucleic acid
CC sequences encoding them, and propeptide sequences. The activity of
CC the peptides of the invention may be described as, analgesic,
CC anticonvulsant, vasotropic, cardiant, neuroprotective, cerebroprotective,
CC cardiovascular, antiinflammatory, antimigraine, antidiabetic,
CC tranquiliser, vulnerary, antipsychotic, anxiolytic and neuroleptic.
CC Peptides of the invention act by modulating the activity of voltage gated
CC ion channels. They may be used for treating or preventing disorders
CC associated with voltage gated ion channels such as neurological
CC disorders, e.g. seizure (associated with epilepsy), neurotoxic injury
CC associated with conditions of hypoxia, anoxia, ischaemia, stroke,
CC cerebrovascular accident, brain or spinal chord trauma, drowning,
CC suffocation, perinatal asphyxia or hypoglycaemic events; pain e.g.
CC migraine; inflammation or cardiovascular disorders. They may also be used
CC for treating psychiatric disorders e.g. psychosis, anxiety or
CC schizophrenia. The analgesic agents of the invention show diminished side
CC effects and toxicity and are non-addictive. The sequences given in
CC records ABB96595-ABB96697 represent omega-conopeptide propeptide
CC sequences.
CCSQ Sequence 74 AA;
SQ

Query Match

Best Local Similarity

Matches

DB

56.0%; Score 46.5; DB 23; Length 74;

56.2%; Pred. No. 29;

0; Mismatches

2; Indels

5; Gaps

1;

1 NGVCCGY-----KLCH 11

55 NRYCCGYCKRTHLCH 70

Search completed: December 2, 2002, 10:07:06

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OW protein - protein search, using sw model

Run on: December 2, 2002, 10:06:06 ; Search time 12 seconds
(without alignments)
31.875 Million cell updates/sec

Title: US-09-787-986A-1
Perfect score: 83
Sequence: 1 NGVCCGYKLCXHC 13

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	51.8	574	1 US-08-049-473-2	Sequence 2, Appli
2	43	51.8	574	1 US-08-312-648-2	Sequence 2, Appli
3	43	51.8	574	5 PCT-US94-04190-2	Sequence 2, Appli
4	42	50.6	646	4 US-09-625-188-10	Sequence 10, Appli
5	41	49.4	171	3 US-08-833-167-44	Sequence 44, Appli
6	41	49.4	171	3 US-08-833-167-45	Sequence 45, Appli
7	41	49.4	171	3 US-08-833-167-47	Sequence 47, Appli
8	41	49.4	171	3 US-08-833-167-57	Sequence 57, Appli
9	41	49.4	171	3 US-08-833-167-59	Sequence 59, Appli
10	41	49.4	171	3 US-08-833-167-60	Sequence 60, Appli
11	41	49.4	171	4 US-09-344-837A-44	Sequence 44, Appli
12	41	49.4	171	4 US-09-344-837A-45	Sequence 45, Appli
13	41	49.4	171	4 US-09-344-837A-47	Sequence 47, Appli
14	41	49.4	171	4 US-09-344-837A-57	Sequence 57, Appli
15	41	49.4	171	4 US-09-344-837A-59	Sequence 59, Appli
16	41	49.4	171	4 US-09-344-837A-60	Sequence 60, Appli
17	41	49.4	173	1 US-08-354-456A-7	Sequence 7, Appli
18	41	49.4	174	1 US-08-010-099-82	Sequence 82, Appli
19	41	49.4	174	1 US-08-225-224-5	Sequence 5, Appli
20	41	49.4	174	1 US-08-434-411-2	Sequence 2, Appli
21	41	49.4	174	1 US-08-434-402-2	Sequence 2, Appli
22	41	49.4	174	1 US-08-448-716-82	Sequence 82, Appli
23	41	49.4	174	1 US-08-783-288-2	Sequence 31, Appli
24	41	49.4	174	2 US-08-431-459A-31	Sequence 31, Appli
25	41	49.4	174	2 US-08-890-640-2	Sequence 3, Appli
26	41	49.4	174	2 US-08-722-258-5	Sequence 5, Appli
27	41	49.4	174	3 US-08-833-167-49	Sequence 49, Appli

28	41	49.4	174	3 US-08-833-167-50	Sequence 50, Appli
29	41	49.4	174	3 US-08-833-167-51	Sequence 51, Appli
30	41	49.4	174	3 US-08-833-167-52	Sequence 52, Appli
31	41	49.4	174	3 US-08-833-167-53	Sequence 53, Appli
32	41	49.4	174	3 US-08-833-167-54	Sequence 54, Appli
33	41	49.4	174	3 US-08-833-167-55	Sequence 55, Appli
34	41	49.4	174	3 US-08-833-167-56	Sequence 56, Appli
35	41	49.4	174	3 US-08-833-167-95	Sequence 95, Appli
36	41	49.4	174	3 US-08-833-167-96	Sequence 96, Appli
37	41	49.4	174	3 US-08-833-167-97	Sequence 97, Appli
38	41	49.4	174	3 US-08-833-167-98	Sequence 98, Appli
39	41	49.4	174	3 US-08-833-167-99	Sequence 99, Appli
40	41	49.4	174	3 US-08-833-167-100	Sequence 100, App
41	41	49.4	174	3 US-08-833-167-101	Sequence 101, App
42	41	49.4	174	3 US-08-833-167-102	Sequence 102, App
43	41	49.4	174	3 US-08-833-167-103	Sequence 103, App
44	41	49.4	174	4 US-09-221-181-1	Sequence 1, Appli
45	41	49.4	174	4 US-09-304-186-82	Sequence 82, Appli

ALIGNMENTS

RESULT 1
US-08-049-473-2
: Sequence 2, Application US/08049473
: Patent No. 5386021
: GENERAL INFORMATION:
: APPLICANT: Moss, Joel
: APPLICANT: Mishima, Koichi
: APPLICANT: Nishigangale, Maria
: APPLICANT: Tsuchiya, Mikako
: TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
: TITLE OF INVENTION: PROTEIN WITH AN ADP-RIBOSYLATION FACTOR DOMAIN
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: KNOBE, MARTENS, OLSON AND BEAR
: STREET: 620 NEWPORT CENTER DRIVE
: CITY: NEWPORT BEACH
: STATE: CA
: COUNTRY: USA
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/049,473
: FILING DATE: 19930419
: CLASSIFICATION: 436
: ATTORNEY/AGENT INFORMATION:
: NAME: Fuller, Michael L.
: REGISTRATION NUMBER: 36,516
: REFERENCE/DOCKET NUMBER: NIH050.001CPL
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-235-8550
: TELEFAX: 619-235-0176
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 574 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-049-473-2
Query Match 51.8%; Score 43; DB 1; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 16; Gaps 1;
OY 2 GVC-----CGYKLCXHC 13
||| :|||
DB 32 GVCEDVFSLGDKVPRLLICGHTVCCHDC 59

```
RESULT 2
US-08-312-648-2
; Sequence 2, Application US/08312648
; Patent No. 5514600
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Mishima, Koichi
; APPLICANT: Nightlingale, Maria
; APPLICANT: Tsuchiya, Mikako
; TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
; TITLE OF INVENTION: PROTEIN WITH AN ADP-RIBOSYLATION FACTOR DOMAIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,648
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/049,473
; FILING DATE: 19-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael L.
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH050.001DVI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 574 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-312-648-2

Query Match 51.8%; Score 43; DB 1; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 16; Gaps 1;

QY 2 GVC-----CGKLCXHC 13
|||
|||:||||
Db 32 GVCEDVFSLQDGVRLLCGHTVCHDC 59

RESULT 3
PCT-US94-04190-2
; Sequence 2, Application PC/TUS9404190
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary, Department
; APPLICANT: of Health and Human Services
; TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
; TITLE OF INVENTION: PROTEIN WITH AN ADP-RIBOSYLATION FACTOR DOMAIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
```

```
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04190
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael L.
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH050.001QPC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 574 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-04190-2

Query Match 51.8%; Score 43; DB 5; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 16; Gaps 1;

QY 2 GVC-----CGKLCXHC 13
|||
|||:||||
Db 32 GVCEDVFSLQDGVRLLCGHTVCHDC 59

RESULT 4
US-09-625-188-10
; Sequence 10, Application US/09625188
; Patent No. 6307037
; GENERAL INFORMATION:
; APPLICANT: No. 6307037artis AG
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-31285PI
; CURRENT APPLICATION NUMBER: US/09/625,188
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Ashbya gossypii
US-09-625-188-10

Query Match 50.6%; Score 42; DB 4; Length 646;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 CGKLCXHC 13
||||:|
Db 52 CGYQICQFC 60

RESULT 5
US-08-833-167-44
; Sequence 44, Application US/08833167
; Patent No. 6100070
; GENERAL INFORMATION:
; APPLICANT: ZURFLUH, LINDA L
; APPLICANT: MCHESTER, CHARLES A
; APPLICANT: MCKEARN, JOHN P
; APPLICANT: KLEIN, BARBARA K
; APPLICANT: PENG, YIQUING
; APPLICANT: BRAFORD-GOLDBERG, SARAH R
; APPLICANT: LEE, STEPHEN C
```

;; TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
;; NUMBER OF SEQUENCES: 129
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
;; ADDRESSEE: CORPORATE PATENT DEPT.
;; STREET: P.O. BOX 5110
;; CITY: CHICAGO
;; STATE: ILLINOIS
;; COUNTRY: USA
;; ZIP: 60680
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/833,167
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US 96/15935
;; FILING DATE: 04-OCT-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/004,382
;; FILING DATE: 05-OCT-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENNETT, DENNIS A
;; REFERENCE/DOCKET NUMBER: 2907/1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 314-694-5402
;; TELEFAX: 314-694-9095
;; INFORMATION FOR SEQ ID NO: 44:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 171 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-833-167-44
;;
Query Match 49.4%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 VCCGYKLCH 11
: | | | | |
DB 110 LCATYKLCH 118

RESULT 6
US-08-833-167-45
; Sequence 45, Application US/08833167
; Patent No. 6100070
; GENERAL INFORMATION:
; APPLICANT: ZURFLUH, LINDA L
; APPLICANT: MCMHERTER, CHARLES A
; APPLICANT: MCKEARN, JOHN P
; APPLICANT: KLEIN, BARBARA K
; APPLICANT: FENG, YI QING
; APPLICANT: BRAFORD-GOLDBERG, SARAH R
; APPLICANT: LEE, STEPHEN C
; TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
; NUMBER OF SEQUENCES: 129
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
; ADDRESSEE: CORPORATE PATENT DEPT.
; STREET: P.O. BOX 5110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/833,167
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US 96/15935
;; FILING DATE: 04-OCT-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/004,382
;; FILING DATE: 05-OCT-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENNETT, DENNIS A
;; REFERENCE/DOCKET NUMBER: 2907/1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 314-694-5402
;; TELEFAX: 314-694-9095
;; INFORMATION FOR SEQ ID NO: 45:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 171 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-833-167-45
;;
Query Match 49.4%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 VCCGYKLCH 11
: | | | | |
DB 74 LCATYKLCH 82

RESULT 7
US-08-833-167-47
; Sequence 47, Application US/08833167
; Patent No. 6100070
; GENERAL INFORMATION:
; APPLICANT: ZURFLUH, LINDA L
; APPLICANT: MCMHERTER, CHARLES A
; APPLICANT: MCKEARN, JOHN P
; APPLICANT: KLEIN, BARBARA K
; APPLICANT: FENG, YI QING
; APPLICANT: BRAFORD-GOLDBERG, SARAH R
; APPLICANT: LEE, STEPHEN C
; TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
; NUMBER OF SEQUENCES: 129
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
; ADDRESSEE: CORPORATE PATENT DEPT.
; STREET: P.O. BOX 5110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,167
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US 96/15935
; FILING DATE: 04-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,382

FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENNETT, DENNIS A
REFERENCE/DOCKET NUMBER: 2907/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-5402
TELEFAX: 314-694-9095
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-167-47

Query Match 49.4%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCH 11
: | |||||
DB 65 LCATYKLCH 73

RESULT 8
US-08-833-167-57
Sequence 57, Application US/08833167
Patent No. 6100070
GENERAL INFORMATION:
APPLICANT: ZURFLUH, LINDA L
APPLICANT: MCHESTER, CHARLES A
APPLICANT: MCKEARN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: PENG, YIOWING
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,167
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENNETT, DENNIS A
REFERENCE/DOCKET NUMBER: 2907/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-5402
TELEFAX: 314-694-9095
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-167-57

Query Match 49.4%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCH 11
: | |||||
DB 110 LCATYKLCH 118

RESULT 9
US-08-833-167-59
Sequence 59, Application US/08833167
Patent No. 6100070
GENERAL INFORMATION:
APPLICANT: ZURFLUH, LINDA L
APPLICANT: MCHESTER, CHARLES A
APPLICANT: MCKEARN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: PENG, YIOWING
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,167
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENNETT, DENNIS A
REFERENCE/DOCKET NUMBER: 2907/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-5402
TELEFAX: 314-694-9095
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-167-59

Query Match 49.4%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCH 11
: | |||||
DB 74 LCATYKLCH 82

RESULT 10
US-08-833-167-60
Sequence 60, Application US/08833167
Patent No. 6100070
GENERAL INFORMATION:
APPLICANT: ZURELUH, LINDA L
APPLICANT: MCWHERTER, CHARLES A
APPLICANT: MCKEARN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIQUING
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
ADDRESSEE: CORPORATE PATENT DEPT.,
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,167
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENNETT, DENNIS A
REFERENCE/DOCKET NUMBER: 2907/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-5402
TELEFAX: 314-694-9095
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-167-60

Query Match 49.4%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKRLCH 11
: 1 11111
Db 65 LCATYKRLCH 73

RESULT 11
US-09-344-837A-44
Sequence 44, Application US/09344837A
Patent No. 6358505
GENERAL INFORMATION:
APPLICANT: ZURELUH, LINDA L
APPLICANT: MCWHERTER, CHARLES A
APPLICANT: MCKEARN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIQUING
APPLICANT: BRAFORD-GOLDBERG, SARAH R

APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: S. CHRISTOPHER BAUER
ADDRESSEE: MONSANTO/G. D. SEARLE & CO.,
ADDRESSEE: PATENT DEPARTMENT CENTRAL
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,837A
FILING DATE: 25-JUN-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: S. CHRISTOPHER BAUER
REFERENCE/DOCKET NUMBER: 2907/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 636-737-6257
TELEFAX: 636-737-5452
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-344-837A-44

Query Match 49.4%; Score 41; DB 4; Length 171;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKRLCH 11
: 1 11111
Db 110 LCATYKRLCH 118

RESULT 12
US-09-344-837A-45
Sequence 45, Application US/09344837A
Patent No. 6358505
GENERAL INFORMATION:
APPLICANT: ZURELUH, LINDA L
APPLICANT: MCWHERTER, CHARLES A
APPLICANT: MCKEARN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIQUING
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: S. CHRISTOPHER BAUER
ADDRESSEE: MONSANTO/G. D. SEARLE & CO.,
ADDRESSEE: PATENT DEPARTMENT CENTRAL
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA

ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,837A
FILING DATE: 25-JUN-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: S. CHRISTOPHER BAUER
REFERENCE/DOCKET NUMBER: 2907/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 636-737-6257
FAX: 636-737-5452
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-344-837A-45

Query Match 49.4%; Score 41; DB 4; Length 171;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 VCCGYKLC 11
: | | | | |
DB 74 LCATYKLC 82

RESULT 13
US-09-344-837A-47
Sequence 47, Application US/09344837A
Patent No. 6358505
GENERAL INFORMATION:
APPLICANT: ZURFLUH, LINDA L.
APPLICANT: MCKERTER, CHARLES A
APPLICANT: MCKERN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIOLING
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: S. CHRISTOPHER BAUER
ADDRESSER: MONSANTO/G. D. SEARLE & CO.,
STREET: P. O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,837A
FILING DATE: 25-JUN-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: S. CHRISTOPHER BAUER
REFERENCE/DOCKET NUMBER: 2907/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 636-737-6257
FAX: 636-737-5452
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-344-837A-47

Query Match 49.4%; Score 41; DB 4; Length 171;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 VCCGYKLC 11
: | | | | |
DB 65 LCATYKLC 73

RESULT 14
US-09-344-837A-57
Sequence 57, Application US/09344837A
Patent No. 6358505
GENERAL INFORMATION:
APPLICANT: ZURFLUH, LINDA L.
APPLICANT: MCKERTER, CHARLES A
APPLICANT: MCKERN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIOLING
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: S. CHRISTOPHER BAUER
ADDRESSER: MONSANTO/G. D. SEARLE & CO.,
STREET: P. O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,837A
FILING DATE: 25-JUN-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: S. CHRISTOPHER BAUER
REFERENCE/DOCKET NUMBER: 2907/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 636-737-6257
FAX: 636-737-5452

; INFORMATION FOR SEQ ID NO: 57:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 171 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-344-837A-57

Query Match 49.4%; Score 41; DB 4; Length 171;
 Best Local Similarity 66.7%; Pred. No. 76;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCH 11
 : | |||||
 Db 110 LCATYKLCH 118

RESULT 15
 US-09-344-837A-59
 ; Sequence 59, Application US/09344837A
 ; Patent No. 6358505

; GENERAL INFORMATION:

; APPLICANT: ZURFLOH, LINDA L
 ; APPLICANT: MCWHERTER, CHARLES A
 ; APPLICANT: MCKEARN, JOHN P
 ; APPLICANT: KLEIN, BARBARA K
 ; APPLICANT: FENG, YIONG
 ; APPLICANT: BRADFORD-GOLDBERG, SARAH R
 ; APPLICANT: LEE, STEPHEN C
 ; TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
 ; NUMBER OF SEQUENCES: 129
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: S. CHRISTOPHER BAUER
 ; ADDRESSEE: MONSANTO/G. D. SEARLE & CO.,
 ; ADDRESSEE: PATENT DEPARTMENT CENTRAL
 ; STREET: P. O. BOX 5110
 ; CITY: CHICAGO
 ; STATE: ILLINOIS
 ; COUNTRY: USA
 ; ZIP: 60680

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/344,837A
 ; FILING DATE: 25-JUN-1999
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/US 96/15935
 ; FILING DATE: 04-OCT-1996
 ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/004,382
 ; FILING DATE: 05-OCT-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: S. CHRISTOPHER BAUER
 ; REFERENCE/DOCKET NUMBER: 2907/2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 636-737-6257
 ; TELEFAX: 636-737-5452
 ; INFORMATION FOR SEQ ID NO: 59:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 171 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-344-837A-59

Query Match 49.4%; Score 41; DB 4; Length 171;
 Best Local Similarity 66.7%; Pred. No. 76;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VCCGYKLCH 11
 : | |||||
 Db 74 LCATYKLCH 82

Search completed: December 2, 2002, 10:09:27
 Job time : 13 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:07:11 ; Search time 7.5 Seconds
(without alignments)
27.602 Million cell updates/sec

Title: US-09-787-986a-1
Perfect score: 83
Sequence: 1 NGVCCGYKLCXHC 13

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCF_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCFOS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	46	55.4	538	10	US-09-779-307-2
2	46	55.4	538	10	US-09-779-307-11
3	46	55.4	547	10	US-09-779-307-12
4	46	55.4	547	10	US-09-779-307-13
5	43	51.8	26	8	US-08-775-765-9
6	42	50.6	24	8	US-08-775-765-8
7	41	49.4	174	10	US-09-754-532-82
8	41	49.4	174	10	US-09-760-008A-1
9	41	49.4	174	10	US-09-950-473-2
10	41	49.4	174	12	US-10-016-403-1
11	41	49.4	174	12	US-10-016-403-3
12	41	49.4	174	12	US-10-016-403-4
13	41	49.4	174	12	US-10-016-403-4
14	41	49.4	174	12	US-10-003-496-1
15	41	49.4	175	10	US-09-754-532-2
16	41	49.4	175	10	US-09-754-532-67
17	41	49.4	175	10	US-09-754-532-68
18	41	49.4	175	10	US-09-754-532-69
19	41	49.4	175	10	US-09-754-532-71

20	41	49.4	175	10	US-09-754-532-76	Sequence 76, Appl
21	41	49.4	175	10	US-09-754-532-78	Sequence 78, Appl
22	41	49.4	175	10	US-09-754-532-79	Sequence 79, Appl
23	41	49.4	175	10	US-09-754-532-80	Sequence 80, Appl
24	41	49.4	175	10	US-09-754-532-81	Sequence 81, Appl
25	41	49.4	175	10	US-09-754-532-85	Sequence 85, Appl
26	41	49.4	175	10	US-09-754-532-86	Sequence 86, Appl
27	41	49.4	175	10	US-09-754-532-87	Sequence 87, Appl
28	41	49.4	175	10	US-09-754-532-88	Sequence 88, Appl
29	41	49.4	175	10	US-09-754-532-89	Sequence 89, Appl
30	41	49.4	175	10	US-09-754-532-90	Sequence 90, Appl
31	41	49.4	175	10	US-09-754-532-91	Sequence 91, Appl
32	41	49.4	175	10	US-09-754-532-92	Sequence 92, Appl
33	41	49.4	175	10	US-09-754-532-93	Sequence 93, Appl
34	41	49.4	175	10	US-09-754-532-94	Sequence 94, Appl
35	41	49.4	175	10	US-09-754-532-95	Sequence 95, Appl
36	41	49.4	175	10	US-09-754-532-96	Sequence 96, Appl
37	41	49.4	175	10	US-09-754-532-97	Sequence 97, Appl
38	41	49.4	175	10	US-09-754-532-98	Sequence 98, Appl
39	41	49.4	175	10	US-09-754-532-99	Sequence 99, Appl
40	41	49.4	175	10	US-09-754-532-100	Sequence 100, App
41	41	49.4	175	10	US-09-754-532-101	Sequence 101, App
42	41	49.4	175	10	US-09-754-532-102	Sequence 102, App
43	41	49.4	175	10	US-09-754-532-104	Sequence 104, App
44	41	49.4	175	10	US-09-754-532-105	Sequence 105, App
45	41	49.4	175	10	US-09-754-532-106	Sequence 106, App

ALIGNMENTS

RESULT 1
US-09-779-307-2
; Sequence 2, Application US/09779307
; Patent No. US20020137675A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corinne
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-662 US
; CURRENT APPLICATION NUMBER: US/09/779,307
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-307-2

Query Match 55.4%; Score 46; DB 10; Length 538;
Best Local Similarity 50.0%; Pred No. 34;
Matches 8; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 2 GVCC---GYKLCXHC 13
1 11 1: 111 1
Db 508 GCCCFPLDGLLCHGC 523

RESULT 2
US-09-779-307-11
; Sequence 11, Application US/09779307
; Patent No. US20020137675A1

```

; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Majumder, Kunud
; APPLICANT: Verneet, Corine
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-662 US
; CURRENT FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 11
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-307-11

```

```

Query Match          55.4%; Score 46; DB 10; Length 538;
Best Local Similarity 50.0%; Pred. No. 34;
Matches      8; Conservative      1; Mismatches      3; Indels      4; Gaps      1;

```

```

QY      2 GVCC----GYKLCHXC 13
      1 11 1: 111 1
Db      508 GCCCFPLDGHLLCHGC 523

```

```

RESULT 3
US-09-779-307-12
; Sequence 12, Application US/09779307
; Patent No. US20020137675A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Majumder, Kunud
; APPLICANT: Verneet, Corine
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded thereby
; FILE REFERENCE: 15966-662 US
; CURRENT APPLICATION NUMBER: US/09/779,307
; CURRENT FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 12
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-779-307-12

```

```

Query Match          55.4%; Score 46; DB 10; Length 547;
Best Local Similarity 50.0%; Pred. No. 35;
Matches      8; Conservative      1; Mismatches      3; Indels      4; Gaps      1;

```

```

QY      2 GVCC----GYKLCHXC 13
      1 11 1: 111 1
Db      517 GCCCFPLDGHLLCHGC 532

```

RESULT 4

```

US-09-779-307-13
; Sequence 13, Application US/09779307
; Patent No. US20020137675A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Majumder, Kunud
; APPLICANT: Verneet, Corine
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-662 US
; CURRENT FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 13
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-779-307-13

```

```

Query Match          55.4%; Score 46; DB 10; Length 547;
Best Local Similarity 50.0%; Pred. No. 35;
Matches      8; Conservative      1; Mismatches      3; Indels      4; Gaps      1;

```

```

QY      2 GVCC----GYKLCHXC 13
      1 11 1: 111 1
Db      517 GCCCFPLDGHLLCHGC 532

```

```

RESULT 5
US-08-775-765-9
; Sequence 9, Application US/08775765C
; Patent No. US20010010821A1
; GENERAL INFORMATION:
; APPLICANT: Kelleher, Dermot
; APPLICANT: Windle, Henry
; APPLICANT: Byrne, William
; APPLICANT: McManus, Ross
; TITLE OF INVENTION: Helicobacter Proteins and Vaccines
; FILE REFERENCE: 08/775 765
; CURRENT APPLICATION NUMBER: US/08/775,765C
; CURRENT FILING DATE: 1996-12-31
; EARLIER APPLICATION NUMBER: IE 94 0538
; EARLIER FILING DATE: 1994-07-01
; EARLIER APPLICATION NUMBER: IE 95 0249
; EARLIER FILING DATE: 1995-04-06
; EARLIER APPLICATION NUMBER: PCT/IE 95/00036
; EARLIER FILING DATE: 1995-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 9
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-775-765-9

```

```

Query Match          51.8%; Score 43; DB 8; Length 26;
Best Local Similarity 66.7%; Pred. No. 7.8;
Matches      6; Conservative      1; Mismatches      2; Indels      0; Gaps      0;

```

```

QY      2 GVCCGYKLC 10
      1 1111 1
Db      8 GCCCGYYIC 16

```

RESULT 6
US-08-775-765-8
; Sequence 8, Application US/08775765C
; Patent No. US20010010821A1
; GENERAL INFORMATION:
; APPLICANT: Kelleher, Dermot
; APPLICANT: Windle, Henry
; APPLICANT: Byrne, William
; APPLICANT: McManus, Ross
; TITLE OF INVENTION: Helicobacter Proteins and Vaccines
; FILE REFERENCE: 08/775 765
; CURRENT APPLICATION NUMBER: US/08/775,765C
; CURRENT FILING DATE: 1996-12-31
; EARLIER APPLICATION NUMBER: IE 94 0538
; EARLIER FILING DATE: 1994-07-01
; EARLIER APPLICATION NUMBER: IE 95 0249
; EARLIER FILING DATE: 1995-04-06
; EARLIER APPLICATION NUMBER: PCT/IE 95/00036
; EARLIER FILING DATE: 1995-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-775-765-8

Query Match
Best Local Similarity 50.6%; Score 42; DB 8; Length 24;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 GVCQGYKLCXHC 13
DB 6 GGCCGTYTCYTC 17

RESULT 7
US-09-754-532-82
; Sequence 82, Application US/09754532
; Patent No. US20010016191A1
; GENERAL INFORMATION:
; APPLICANT: Osslund, Timothy D.
; TITLE OF INVENTION: G-CSF ANALOG COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: Amgen Center, 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States of America
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/754,532
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/448,716
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pessin, Karol
; REGISTRATION NUMBER: 34,899
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 805/499-5725
; TELEFAX: 805/499-8011
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-754-532-82

Query Match
Best Local Similarity 49.4%; Score 41; DB 10; Length 174;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCX 11
DB 36 LCATYKLCX 44

RESULT 8
US-09-760-008A-1
; Sequence 1, Application US/09760008A
; Patent No. US2002004483A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
; APPLICANT: MIKKELSEN, JAN MOLLER
; TITLE OF INVENTION: G-CSF COMPOUNDS
; FILE REFERENCE: 31-000700US
; CURRENT APPLICATION NUMBER: US/09/760,008A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: DK PA 2000 00341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-008A-1

Query Match
Best Local Similarity 49.4%; Score 41; DB 10; Length 174;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCX 11
DB 35 LCATYKLCX 43

RESULT 9
US-09-950-473-2
; Sequence 2, Application US/09950473
; Patent No. US20020151488A1
; GENERAL INFORMATION:
; APPLICANT: Sarkar, Casim
; APPLICANT: lauffenburger, Douglas
; TITLE OF INVENTION: G-CSF Analog Compositions and Methods
; FILE REFERENCE: 0101737732
; CURRENT APPLICATION NUMBER: US/09/950,473
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-473-2

Query Match 49.4%; Score 41; DB 10; Length 174;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 VCCGYKLCH 11
: 1 11111
DB 35 LCATYKLCH 43

RESULT 10
US-10-016-403-1
; Sequence 1, Application US/10016403
; Patent No. US20020107505A1
; GENERAL INFORMATION:
; APPLICANT: Holladay, Leslie A.
; TITLE OF INVENTION: MODIFICATION OF POLYPEPTIDE DRUGS TO
; INCREASE ELECTROTRANSPORT FLUX
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stroud, Stroud, Willink, Thompson & Howard
; STREET: 25 West Main Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2236
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/016,403
; FILING DATE: 10-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/466,610
; FILING DATE: 1995-JUN-06
; ATTORNEY/AGENT INFORMATION:
; NAME: Frenchick, Grady J.
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 8734.28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-257-2281
; TELEFAX: 608-257-7643
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..174
; OTHER INFORMATION:
; OTHER INFORMATION: stimulating factor"
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-016-403-1
Query Match 49.4%; Score 41; DB 12; Length 174;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 3 VCCGYKLCH 11
: 1 11111
DB 35 LCATYKLCH 43
RESULT 11
US-10-016-403-2
; Sequence 2, Application US/10016403
; Patent No. US20020107505A1
; GENERAL INFORMATION:
; APPLICANT: Holladay, Leslie A.

; TITLE OF INVENTION: MODIFICATION OF POLYPEPTIDE DRUGS TO
; INCREASE ELECTROTRANSPORT FLUX
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stroud, Stroud, Willink, Thompson & Howard
; STREET: 25 West Main Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2236
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/016,403
; FILING DATE: 10-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/466,610
; FILING DATE: 1995-JUN-06
; ATTORNEY/AGENT INFORMATION:
; NAME: Frenchick, Grady J.
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 8734.28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-257-2281
; TELEFAX: 608-257-7643
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..174
; OTHER INFORMATION:
; OTHER INFORMATION: stimulating factor"
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-016-403-2
Query Match 49.4%; Score 41; DB 12; Length 174;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 3 VCCGYKLCH 11
: 1 11111
DB 35 LCATYKLCH 43
RESULT 12
US-10-016-403-3
; Sequence 3, Application US/10016403
; Patent No. US20020107505A1
; GENERAL INFORMATION:
; APPLICANT: Holladay, Leslie A.
; TITLE OF INVENTION: MODIFICATION OF POLYPEPTIDE DRUGS TO
; INCREASE ELECTROTRANSPORT FLUX
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stroud, Stroud, Willink, Thompson & Howard
; STREET: 25 West Main Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2236
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/016,403
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/466,610
FILING DATE: 1995-JUN-06
ATTORNEY/AGENT INFORMATION:
NAME: Frenchick, Grady J.
REGISTRATION NUMBER: 29,018
REFERENCE/DOCKET NUMBER: 8734,28
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-257-2281
TELEFAX: 608-257-7643
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..174
OTHER INFORMATION: /note= "modified granulocyte-colony
stimulating factor"
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-016-403-3

Query Match 49.4%; Score 41; DB 12; Length 174;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 VCCGKYLCH 11
: 1 11111
DB 35 LCATYKLCH 43

RESULT 13
US-10-016-403-4
Sequence 4, Application US/10016403
Patent No. US20020107505A1
GENERAL INFORMATION:
APPLICANT: Holladay, Leslie A.
TITLE OF INVENTION: MODIFICATION OF POLYPEPTIDE DRUGS TO
INCREASE ELECTROTRANSPORT FLUX
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stroud, Willink, Thompson & Howard
STREET: 25 West Main Street
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2236
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/016,403
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/466,610
FILING DATE: 1995-JUN-06
ATTORNEY/AGENT INFORMATION:
NAME: Frenchick, Grady J.
REGISTRATION NUMBER: 29,018
REFERENCE/DOCKET NUMBER: 8734,28
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-257-2281
TELEFAX: 608-257-7643
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids

TYPE: amino acid
TOPOLOGY: linear
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..174
OTHER INFORMATION: /note= "granulocyte-colony
stimulating factor"
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-016-403-4

Query Match 49.4%; Score 41; DB 12; Length 174;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 VCCGKYLCH 11
: 1 11111
DB 35 LCATYKLCH 43

RESULT 14
US-10-003-496-1
Sequence 1, Application US/10003496
Patent No. US20020142964A1
GENERAL INFORMATION:
APPLICANT: Maxygen Aps
TITLE OF INVENTION: Single-Chain Polypeptides
FILE REFERENCE: 0218us210
CURRENT APPLICATION NUMBER: US/10/003,496
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/245,727
PRIOR FILING DATE: 2000-11-02
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 174
TYPE: PRT
ORGANISM: Homo sapiens
US-10-003-496-1

Query Match 49.4%; Score 41; DB 12; Length 174;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 VCCGKYLCH 11
: 1 11111
DB 35 LCATYKLCH 43

RESULT 15
US-09-754-532-2
Sequence 2, Application US/09754532
Patent No. US20010016191A1
GENERAL INFORMATION:
APPLICANT: Osslund, Timothy D.
TITLE OF INVENTION: G-CSF ANALOG COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: Amgen Center, 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: United States of America
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/754,532
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/448,716
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Pessin, Karol
 REGISTRATION NUMBER: 34,899
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 805/499-5725
 TELEFAX: 805/499-8011
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 175 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-754-532-2

Query Match 49.4%; Score 41; DB 10; Length 175;
 Best Local Similarity 66.7%; Pred. No. 59;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 VCCGYKLC 11
 : | | | | |
 DB 36 LCATYKLC 44

Search completed: December 2, 2002, 10:09:48
 Job time : 8.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:05:26 ; Search time 13.5 Seconds
(without alignments)
92.574 Million cell updates/sec

Title: US-09-787-986a-1

Perfect score: 83
Sequence: 1 NGVCCGYKLCCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	68.7	258	2 T130368	probable immediate
2	48	57.8	1322	2 T15689	hypothetical prote
3	47	56.6	245	2 T10407	immediate early pr
4	47	56.6	322	2 T08592	TGACG-motif bindin
5	47	56.6	326	2 T08591	TGACG-motif bindin
6	47	56.6	373	2 T47488	hypothetical prote
7	46	55.4	55	2 S25774	testis-specific pr
8	45	54.2	79	2 F95210	conserved domain p
9	44	53.0	244	2 T26913	hypothetical prote
10	44	53.0	244	2 T26912	hypothetical prote
11	44	53.0	332	2 T24312	hypothetical prote
12	44	53.0	514	2 T10559	hypothetical prote
13	44	53.0	600	2 T49281	hypothetical prote
14	43	51.8	261	2 G72867	hypothetical prote
15	43	51.8	261	2 T41878	hypothetical prote
16	43	51.8	557	2 T27752	hypothetical prote
17	43	51.8	574	2 A46054	hypothetical prote
18	42	50.6	78	2 T50943	GTP-binding protei
19	42	50.6	104	2 F83370	hydrogen cyanide s
20	42	50.6	126	2 T42321	hypothetical prote
21	42	50.6	211	2 T20590	hypothetical prote
22	42	50.6	222	2 T47487	hypothetical prote
23	42	50.6	350	2 G84647	hypothetical prote
24	42	50.6	373	2 E84647	hypothetical prote
25	42	50.6	400	1 ZBBE14	hypothetical prote
26	42	50.6	489	2 S62474	probable transcrip
27	42	50.6	587	2 A56015	finger protein sig
28	41.5	50.0	1181	2 D86157	hypothetical prote
29	41	49.4	98	2 JC5147	tachyctin precurs

30	41	49.4	171	2 G90687	phosphatidyglycer
31	41	49.4	171	2 C85538	phosphatidyglycer
32	41	49.4	171	2 AD0554	phosphatidyglycer
33	41	49.4	172	2 B64771	phosphatidyglycer
34	41	49.4	194	1 E64381	conserved hypotet
35	41	49.4	204	1 F0HUGL	granulocyte colony
36	41	49.4	204	2 B86308	F20D23.10 protein
37	41	49.4	208	2 A26496	granulocyte colony
38	41	49.4	214	2 JC5043	granulocyte colony
39	41	49.4	295	2 T20629	hypothetical prote
40	41	49.4	348	2 T47494	hypothetical prote
41	41	49.4	372	2 T29359	hypothetical prote
42	41	49.4	884	2 T18649	hypothetical prote
43	41	49.4	1101	2 T16840	hypothetical prote
44	41	49.4	1235	1 QGBEM4	DNA-binding protei
45	40	48.2	179	2 AB3068	hypothetical prote

ALIGNMENTS

RESULT 1

T130368
probable immediate-early transactivator 0 - Lymantria dispar nuclear polyhedrosis vir
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T130368
R:Kuzio, U.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavecek, J.M.; R
Virology 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantri
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T130368
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-258 <KUZ>
A:Cross-references: EMBL:AF081810; PIDN:AAC70206.1
C:Keywords: immediate-early protein

Query Match Score 57; DB 2; Length 258;
Best Local Similarity 63.6%; Pred. No. 0.57;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VCCGYKLCCHXC 13

DB 208 VCCGYKLCCHXC 218

RESULT 2

T15689
hypothetical protein C28G1.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C:Accession: T15689
R:Favell, T.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C28G1.
A:Reference number: Z18389

A:Accession: T15689
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1322 <FAV>

A:Cross-references: EMBL:U41026; NID:g1086701; PID:g1086702; PIDN:AA82350.1; CESP:C2
C:genetics:
A:Gene: CESP:C28G1.3
A:Introns: 25/1; 131/3; 150/1; 166/3; 180/3; 204/2; 235/3; 344/3; 385/2; 436/3; 483/2
C:Superfamily: RING finger homology
F:810-862/Domain: RING finger homology <RRN>

Query Match Score 48; DB 2; Length 1322;
Best Local Similarity 63.6%; Pred. No. 32;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VCCGYKLCCHXC 13

Db 1071 VTGHALCHXC 1081

RESULT 3

T10407
Immediate early protein 0 - Orgyia pseudotsugata nuclear polyhedrosis virus
C:Species: Orgyia pseudotsugata nuclear polyhedrosis virus, OpNPV
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C:Accession: T10407
R:Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohmann, G.F.
Virology 229, 381-399, 1997
A:Title: The sequence of the Orgyia pseudotsugata multinuclеocapsid nuclear polyhedrosis
A:Reference number: 217011; MUID:97271300; PMID:9126251
A:Accession: T10407
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-245 <AHP>
A:Cross-references: EMBL:U75930; NID:g2934903; PID:g1911384

Query Match

Best Local Similarity 56.6%; Score 47; DB 2; Length 245;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 CGYKLCXHC 13
||| ||| |
Db 213 CCYSLCYAC 222

RESULT 4

T08592
TGACG-motif-binding protein STF2 - soybean
C:Species: Glycine max (soybean)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08592
R:Cheong, Y.H.; Yoo, C.M.; Park, J.M.; Ryu, G.R.; Goeckjlan, V.H.; Nagao, R.T.; Key, J.L.
submitted to the EMBL Data Library, September 1995
A:Description: STF1 is a novel TGACG-binding factor with a zinc-finger motif and a bzip
A:Reference number: 216445
A:Accession: T08592
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-322 <CHE>
A:Cross-references: EMBL:L28004; NID:g986966; PID:g2934885
A:Experimental source: strain Williams; hypocotyl
A:Genetics:
A:Gene: STF2

Query Match

Best Local Similarity 56.6%; Score 47; DB 2; Length 322;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
||| ||| |
Db 61 CGPPLCHSC 69

RESULT 5

T08591
TGACG-motif binding protein STF1 - soybean
C:Species: Glycine max (soybean)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08591
R:Cheong, Y.H.; Yoo, C.M.; Park, J.M.; Ryu, G.R.; Goeckjlan, V.H.; Nagao, R.T.; Key, J.L.
submitted to the EMBL Data Library, September 1995
A:Description: STF1 is a novel TGACG-binding factor with a zinc-finger motif and a bzip
A:Reference number: 216445
A:Accession: T08591
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-326 <CHE>
A:Cross-references: EMBL:L28003; NID:g2934883; PID:g2934884
A:Experimental source: strain Williams; hypocotyl

Query Match
Best Local Similarity 56.6%; Score 47; DB 2; Length 326;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
||| ||| |
Db 61 CGPPLCHSC 69

RESULT 6

T47488
Hypothetical protein F9K21.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47488
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, February 2000
A:Reference number: 224467
A:Accession: T47488
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-373 <IOR>
A:Cross-references: EMBL:AL138657
A:Experimental source: cultivar Columbia; BAC clone F9K21
C:Genetics:
A:Map position: 3
A:Introns: 92/3; 328/2; 353/3
A:Note: F9K21.60

Query Match

Best Local Similarity 56.6%; Score 47; DB 2; Length 373;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
||| ||| |
Db 329 CGYKFCYAC 337

RESULT 7

S25774
testis-specific protein Mst84DC - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
C:Accession: S25774; C56565
R:Kuhn, R.; Kuhn, C.; Boersch, D.; Glaetzer, K.H.; Schaefer, U.; Schaefer, M.
Mech. Dev. 35, 143-151, 1991
A:Title: A cluster of four genes selectively expressed in the male germ line of Dros
A:Reference number: A56565; MUID:92102953; PMID:1684716
A:Accession: S25774
A:Molecule type: DNA
A:Residues: 1-55 <RUH>
A:Cross-references: EMBL:X67703; NID:g11072; PID:CAA47939.1; PID:g11075
A:Note: the authors translated the codon TGC for residue 55 as Thr
A:Note: sequence extracted from NCBI backbone (NCBIN:74217, NCBI:P.74222)
C:Genetics:
A:Gene: Mst84DC
A:Cross-references: FlyBase:FBgn0004174
A:Map position: 3
C:Superfamily: fruit fly testis-specific protein
C:Keywords: spermatogenesis; tandem repeat

Query Match

Best Local Similarity 55.4%; Score 46; DB 2; Length 55;
Best Local Similarity 58.3%; Pred. No. 6.4;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 GVCGYKLCXHC 13
||| ||| |
Db 7 GSCGGYCCGPC 18

RESULT 8

F95210
F95210

conserved domain protein SP1806 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: F95210
R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: F95210
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-79 <KUR>
A:Cross-references: GB:AF005672; PIDN:AAK75879.1; PID:q14973305; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP1806

Query Match 54.2%; Score 45; DB 2; Length 79;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CGGKRLCH 11
||| |
DB 41 VCLGKRCH 49

RESULT 9
T26913
hypothetical protein Y45F10B.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26913
R:McMurray, A.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z20286
A:Accession: T26913
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-244 <WIL>
A:Cross-references: EMBL:AL021487; PIDN:CAA16351.1; GSPDB:GN00022; CESP:Y45F10B.8
A:Experimental source: clone Y45F10B
C:Genetics:
A:Gene: CESP:Y45F10B.8
A:Map position: 4
A:Introns: 90/2; 194/1

Query Match 53.0%; Score 44; DB 2; Length 244;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXC 13
||: || |
DB 195 CGHTCHTC 203

RESULT 10
T26912
hypothetical protein Y45F10B.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26912
R:McMurray, A.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z20286
A:Accession: T26912
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-244 <WIL>
A:Cross-references: EMBL:AL021487; PIDN:CAA16350.1; GSPDB:GN00022; CESP:Y45F10B.9
A:Experimental source: clone Y45F10B

C:Genetics:
A:Gene: CESP:Y45F10B.9
A:Map position: 4
A:Introns: 90/2; 194/1

Query Match 53.0%; Score 44; DB 2; Length 244;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXC 13
||: || |
DB 195 CGHTCHTC 203

RESULT 11
T24312
hypothetical protein T01G5.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24312
R:Baslam, V.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19873
A:Accession: T24312
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-332 <WIL>
A:Cross-references: EMBL:Z81111; PIDN:CAB03269.1; GSPDB:GN00023; CESP:T01G5.7
A:Experimental source: clone T01G5
C:Genetics:
A:Gene: CESP:T01G5.7
A:Map position: 5
A:Introns: 208/1; 257/1

Query Match 53.0%; Score 44; DB 2; Length 332;
Best Local Similarity 55.6%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXC 13
||: || |
DB 209 CGHTCHTC 217

RESULT 12
T10559
hypothetical protein F25E4.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
C:Accession: T10559
R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Ban
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16533
A:Accession: T10559
A:Molecule type: DNA
A:Residues: 1-514 <BEV>
A:Cross-references: EMBL:AL050399; GSPDB:GN00062; ATSP:F25E4.10
A:Experimental source: cultivar Columbia; BAC clone F25E4
C:Genetics:
A:Gene: ATSP:F25E4.10
A:Map position: 4
A:Introns: 436/3; 457/3; 479/3

Query Match 53.0%; Score 44; DB 2; Length 514;
Best Local Similarity 60.0%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 CCGYKLCXC 13
|||: || |
DB 295 CCGFYLCFCC 304

RESULT 13
T49281

fertillin alpha precursor - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jun-2000
C:Accession: I49281
R:Wolfsberg, T.G.; Straights, P.D.; Gereña, R.L.; Huovila, A.P.; Primakoff, P.; Myles, D.
Dev. Biol. 169, 378-383, 1995
A:Title: ADAM, a widely distributed and developmentally regulated gene family encoding m
A:Reference number: I48100; MUID:95269891; PMID:7750654
A:Accession: I49281
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-600 <RES>
A:Cross-references: EMBL:U22056; NID:9965009; PIDN:AAAY4920.1; PID:9965010
C:Genetics:
A:Gene: ADAM 1
C:Superfamily: unassigned disintegrins; disintegrin homology
F:246-326/Domain: disintegrin homology <Dis>
F:180/Active site: Glu #status predicted

Query Match 53.0%; Score 44; DB 2; Length 600;
Best Local Similarity 63.6%; Pred. No. 65;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGCCGKRLCH 11
||| :||
DB 485 NGVCNNKRLCH 495

RESULT 14
G72867
hypothetical protein - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
A:Note: dsDNA virus
C>Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
C:Accession: G72867
R:Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A:Reference number: A72850; MUID:94303173; PMID:8030224
A:Accession: G72867
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-261 <AYR>
A:Cross-references: GB:L22858; NID:9510708; PIDN:AAA66771.1; PID:9559210
C:Genetics:
A:Gene: Ac-IE-0

Query Match 51.8%; Score 43; DB 2; Length 261;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCGYKRLCHXC 13
|| :||
DB 230 CCEYAICNAC 239

RESULT 15
T41878
IE-0 orf141 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C:Species: Bombyx mori nuclear polyhedrosis virus, BmSNPV
A:Variety: isolate T3
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C:Accession: T41878
R:Gomi, S.; Matjima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: Z22020; MUID:99281911; PMID:10355780
A:Accession: T41878
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-261 <KAM>
A:Cross-references: EMBL:L33180; PIDN:AAC63807.1
A:Experimental source: isolate T3

C:Genetics:
A:Note: ie-0
Query Match 51.8%; Score 43; DB 2; Length 261;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCGYKRLCHXC 13
|| :||
DB 230 CCEYAICNAC 239

Search completed: December 2, 2002, 10:08:57
Job time : 14.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:04:20 ; Search time 7.5 Seconds
(without alignments)
71.892 Million cell updates/sec

Title: US-09-787-986a-1
Perfect score: 83
Sequence: 1 NGVCCGKLCXHC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	96.4	61	1	P58808 conus marmo
2	74	89.2	13	1	P58810 conus marmo
3	68	81.9	11	1	P58807 conus marmo
4	50	60.2	12	1	P58809 conus marmo
5	47	56.6	245	1	O10369 oryza pseu
6	46	55.4	55	1	Q01644 drosophila
7	43	51.8	261	1	IE0_NPVAC
8	43	51.8	554	1	ARD1_HUMAN
9	42	50.6	400	1	VG78_HSV1
10	42	50.6	400	1	YAC4_SCHRO
11	42	50.6	587	1	NOT4_YEAST
12	42	50.6	837	1	UBPA_DICDI
13	42	50.6	172	1	PGPA_ECOLI
14	41	49.4	194	1	Y653_METJA
15	41	49.4	208	1	CSF3_MOUSE
16	41	49.4	465	1	RN15_HUMAN
17	41	49.4	1235	1	DNBI_HCVNA
18	41	48.2	207	1	CSF3_HUMAN
19	40	48.2	321	1	ABNA_ASPNG
20	40	48.2	469	1	AR11_MOUSE
21	40	48.2	557	1	AR11_HUMAN
22	40	48.2	557	1	PK1_YEAST
23	40	48.2	560	1	PK1_YEAST
24	40	48.2	746	1	CIC5_HUMAN
25	40	48.2	746	1	CIC5_MOUSE
26	40	48.2	746	1	CIC5_MOUSE
27	39.5	47.6	768	1	ITB8_RABIT
28	39.5	47.6	1426	1	EGFR_DROME
29	39	47.0	144	1	NEU2_CAVRO
30	39	47.0	161	1	NEUV_CHICK
31	39	47.0	218	1	VT44_CAEEL
32	39	47.0	257	1	RRBF_SALTY
33	39	47.0	973	1	RRPO_PAV

34	39	47.0	1191	1	DNBI_MCAYS	P30672 murine cyto
35	39	47.0	4486	1	DYH9_HUMAN	Q9nyvc homo sapien
36	38.5	46.4	1203	1	MGR5_RAT	P31424 rattus norv
37	38.5	46.4	1212	1	MGR5_HUMAN	P41594 homo sapien
38	38	45.8	72	1	YVAV_VACCC	P20530 vaccinia vi
39	38	45.8	125	1	NEUM_BUEJA	P08162 buto japoni
40	38	45.8	155	1	NEU4_CATCO	P16229 catostomus
41	38	45.8	158	1	VE6_HPVA5	P21735 human papil
42	38	45.8	194	1	CSF3_FELCA	O02708 felis silve
43	38	45.8	195	1	CSF3_BOVIN	P35833 bos taurus
44	38	45.8	224	1	IAPL_ASFB7	O65138 african swi
45	38	45.8	224	1	IAPL_ASFC3	O11451 african swi

ALIGNMENTS

RESULT 1	ID	Sequence	Standard	PRT	61 AA
AC	P58808	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Lambda-conotoxin CM-VIB precursor (Chi-conotoxin M1A) (Chi-M1A)				
DE	(mrt10a)				
OS	Conus marmoreus (Marble cone).				
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;				
OC	Neogastropoda; Conoidea; Conidae; Conus.				
OX	NCBI_TaxID=42752;				
RP	SEQUENCE FROM N.A., SEQUENCE OF 49-61, SYNTHESIS, AND MASS SPECTROMETRY.				
RP	SEQUENCE FROM N.A., SEQUENCE OF 49-61, SYNTHESIS, AND MASS SPECTROMETRY.				
RC	TISSUE=Venom duct, and Venom;				
RC	TISSUE=Venom; PubMed=10900201;				
RA	Medline=20490660; PubMed=10900201;				
RA	McIntosh J.M., Corpuz G.O., Layer R.T., Garrett J.E., Wagstaff J.D.,				
RA	Bulaj G., Ymazovkina A., Yoshikami D., Cruz L.J., Olivera B.M.;				
RT	"Isolation and characterization of a novel conus peptide with apparent				
RT	antihypertensive activity."				
RL	J. Biol. Chem. 275:32391-32397(2000).				
RL	[2]				
RN	SEQUENCE OF 49-61, AND MASS SPECTROMETRY.				
RP	TISSUE=Venom;				
RC	Medline=20564325; PubMed=10988292;				
RA	Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,				
RA	Seow K.T., Bay B.-H.;				
RT	"Lambda-conotoxins, a new family of conotoxins with unique disulfide				
RT	pattern and protein folding. Isolation and characterization from the				
RT	venom of Conus marmoreus."				
RL	J. Biol. Chem. 275:39516-39522(2000).				
RN	[3]				
RP	SEQUENCE OF 49-61, SYNTHESIS, MASS SPECTROMETRY, AND STRUCTURE BY NMR.				
RC	TISSUE=Venom;				
RC	Medline=21419681; PubMed=11528421;				
RA	Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,				
RA	Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,				
RA	Lewis R.J.;				
RT	"Two new classes of conopeptides inhibit the alpha1-adrenoceptor and				
RT	noradrenergic transporter."				
RL	Nat. Neurosci. 4:902-907(2001).				
CC	- FUNCTION: Inhibits the neuronal noradrenergic transporter.				
CC	- SUBCELLULAR LOCATION: Secreted.				
CC	- TISSUE SPECIFICITY: Expressed by the venom duct.				
CC	- PTM: Exists in two forms, due to cis-trans isomerization at His-				
CC	59-Hyp-60.				
CC	- MASS SPECTROMETRY: MW=1408.5; METHOD=Electrospray.				
CC	- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.				
KW	Neurotoxin; Toxin; Hydroxylation; Signal.				
FT	SIGNAL	1	19	POTENTIAL.	
FT	PROPEP	20	48		
FT	PEPTIDE	49	61	LAMBDA-CONOTOXIN CMVIB.	
FT	DISULFID	52	61		
FT	DISULFID	53	58		

FT MOD_RES 60 60 HYDROXYLATION.
SQ SEQUENCE 61 AA: 6499 MW: F4DE5B5A97EB8DBA CRC64;
Query Match 96.4%; Score 80; DB 1; Length 61;
Best Local Similarity 92.3%; Pred. No. 2.9e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 NGVCGYKLCCHXC 13
Db 49 NGVCGYKLCCHPC 61

RESULT 2
CXL4_CONMR STANDARD: PRT: 13 AA.
ID CXL4_CONMR
AC P58810;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Lambda/chi-conotoxin MrIB (Chi-MrIB).
OS Conus marmoreus (Marble cone).
CC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
CC Neogastropoda; Conoidea; Conus.
OC Neogastropoda; Conoidea; Conus.
OX NCBI_TaxID=42752;
RN (1)
RP SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.
RC TISSUE-Venom;
RX MEDLINE=21419681; PubMed=11528421;
RA Sharpe I.A., Gehmman J., Loughnan M.L., Thomas L., Adams D.A.,
RA Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,
RA Lewis R.J.;
RT "Two new classes of conopeptides inhibit the alpha1-adrenoceptor and
noradrenergic transporter";
RL Nat. Neurosci. 4:902-907(2001).
CC -1 FUNCTION: Inhibits the neuronal noradrenergic transporter.
CC -1 SUBCELLULAR LOCATION: Secreted.
CC -1 TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1 PTM: Exists in two forms, due to cis-trans isomerization at His-
11-Hyp-12.
CC -1 MASS SPECTROMETRY: MW=1393.52; METHOD=Electrospray.
CC -1 SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
DR PDB: 1IEO; 03-APR-02.
KW Neurotoxin; Toxin; Hydroxylation; 3D-structure.
FT DISULFID 4 13
FT MOD_RES 5 10
FT MOD_RES 12 12 HYDROXYLATION.
SQ SEQUENCE 13 AA: 1382 MW: 277AAC376EAD2B58 CRC64;
Query Match 89.2%; Score 74; DB 1; Length 13;
Best Local Similarity 91.7%; Pred. No. 6.4e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 GVCGRYKLCCHXC 13
Db 2 GVCGRYKLCCHPC 13

RESULT 3
CXL1_CONMR STANDARD: PRT: 11 AA.
ID CXL1_CONMR
AC P58807;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Lambda-conotoxin CM-VIA.
OS Conus marmoreus (Marble cone).
CC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
CC Neogastropoda; Conoidea; Conus.
OX NCBI_TaxID=42752;
RN (1)
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE-Venom;
RX MEDLINE=20564325; PubMed=10988292;

RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
pattern and protein folding. Isolation and characterization from the
venom of Conus marmoreus";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -1 FUNCTION: Inhibits the neuronal noradrenergic transporter.
CC -1 SUBCELLULAR LOCATION: Secreted.
CC -1 TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1 MASS SPECTROMETRY: MW=1237.93; MW_ERR=0.21; METHOD=Electrospray.
CC -1 SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 2 11
FT DISULFID 3 8
FT MOD_RES 10 10 HYDROXYLATION.
SQ SEQUENCE 11 AA: 1226 MW: 277AAC60B7232B58 CRC64;
Query Match 81.9%; Score 68; DB 1; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0004;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 3 VCCGRYKLCCHXC 13
Db 1 VCCGRYKLCCHPC 11

RESULT 4
CXL3_CONMR STANDARD: PRT: 12 AA.
ID CXL3_CONMR
AC P58809;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Lambda-conotoxin CM-X.
OS Conus marmoreus (Marble cone).
CC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
CC Neogastropoda; Conoidea; Conus.
OX NCBI_TaxID=42752;
RN (1)
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE-Venom;
RX MEDLINE=20564325; PubMed=10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
pattern and protein folding. Isolation and characterization from the
venom of Conus marmoreus";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -1 FUNCTION: Inhibits the neuronal noradrenergic transporter.
CC -1 SUBCELLULAR LOCATION: Secreted.
CC -1 TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1 MASS SPECTROMETRY: MW=1262.77; MW_ERR=0.07; METHOD=Electrospray.
CC -1 SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 3 12
FT DISULFID 4 9
FT MOD_RES 11 11 HYDROXYLATION.
SQ SEQUENCE 12 AA: 1251 MW: 277AAE242D5A2C8 CRC64;
Query Match 60.2%; Score 50; DB 1; Length 12;
Best Local Similarity 50.0%; Pred. No. 0.15;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 2 GVCGRYKLCCHXC 13
Db 1 GVCGRYKLCCHPC 12

RESULT 5
IEO_NPVOP STANDARD: PRT: 245 AA.
ID IEO_NPVOP
AC O10369;
DT 01-NOV-1997 (Rel. 35, Created)


```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN Immediate-early protein IE-0.
OS Orygia pseudotsugata multicapsid polyhedrosis virus (OpMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolydnavirus.
OX NCBI_TaxID=164623;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohrmann G.F.;
RT "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RT polyhedrosis virus genome."
RL Virology 229:381-399(1997).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U75930; AAC59137.1; -
DR InterPro: IPR001841; Znf_ring
DR PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE: PS50089; ZF_RING_2; 1.
KW Early protein; zinc-finger.
FT ZN-FING 195 240 RING-TYPE.
SQ SEQUENCE 245 AA; 27117 MW; 357FA3B15F7B1029 CRC64;

Query Match 56.6%; Score 47; DB 1; Length 245;
Best Local Similarity 60.0%; Pred. No. 3.7;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
DB 213 CCGYSLCYAC 222

RESULT 6
M84C_DROME STANDARD: PRT: 55 AA.
AC 001644; 09V1A0;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Male specific sperm protein Mst84C.
GN M5F84DC OR CG17945.
OS Drosophila melanogaster (Fruit fly).
OC Euarystola: Metazoa: Arthropoda: Mandibulata: Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Oregon-R;
RX MEDLINE=92102953; PubMed=1684716;
RA Kuhn R., Kuhn C., Boersch D., Glaetzer K.H., Schaefer U.,
RA Schaefer M.;
RT "A cluster of four genes selectively expressed in the male germ line
RT of Drosophila melanogaster."
RL Mech. Dev. 35:143-151(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

```

```

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abdayan A., An H.-J., Andrews-Piankoff C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.-N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.
CC -1- DOMAIN: THIS PROTEIN IS MOSTLY COMPOSED OF REPETITIVE C-G-P
CC MOTIFS.
CC -----
CC -1- SIMILARITY: BELONGS TO THE MST(3)CCP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67703; CAA47939.1; -
DR EMBL: AE003672; AAP54025.1; -
DR FLYBase: FBgn0004174; Mst84C.
KW Spermatogenesis; Repeat; Multigene family.
SQ SEQUENCE 55 AA; 5225 MW; 95A12F3AEC88BD6C CRC64;

Query Match 55.4%; Score 46; DB 1; Length 55;
Best Local Similarity 58.3%; Pred. No. 1.7;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 GVCYKLCXHC 13
DB 7 GSCCYCCGCP 18

RESULT 7
IEO_NPVAC STANDARD: PRT: 261 AA.
AC P41710;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN Immediate-early protein IE-0.

```

```
OS Autographa californica nuclear polyhedrosis virus (ACMPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedroviruses.
OX NCBI_TaxID=46015;
RN SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE=94303173; PubMed=8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
polyhedrosis virus."
RL Virology 202;586-605(1994).
-1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L22858; AAA66771.1; -.
DR InterPro: IPR001841; Znf_finger.
DR PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE: PS50089; ZF_RING_2; 1.
KW Early Protein; Zinc-finger.
FT ZN_FING 212 257
FT SEQUENCE 261 AA; 30109 MW; 7721E0C528EC2CBE CRC64;
SQ
Query Match 51.8%; Score 43; DB 1; Length 261;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 4 CCGYKLCXHC 13
DB 230 CCEYACNAC 239
RESULT 8
ARDL_RAT
ID ARDL_RAT STANDARD: PRT; 554 AA.
AC P36407;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP-binding protein Arp-1 (Fragment).
GN ARPD1 OR ARD1 OR ARP-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93232038; PubMed=8473324;
RA Mishima K., Tsuchiyama M., Nightingale M.S., Moss J., Vaughan M.;
RT "ARD 1, a 64-kDa guanine nucleotide-binding protein with a carboxyl-
terminal ADP-ribosylation factor domain."
RL J. Biol. Chem. 268:8801-8807(1993).
-1- FUNCTION: NOT KNOWN. THE C-TERMINUS CAN ACT AS AN ALLOSTERIC
ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ARP FAMILY
OF GTP-BINDING PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -1- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
```

```
CC -----
DR EMBL: L04760; AAA41301.1; -.
DR HSSP: P32889; IRRG.
DR InterPro: IPR000251; ARF family.
DR InterPro: IPR003649; Bbox_C.
DR InterPro: IPR005225; Small_GTP.
DR InterPro: IPR000315; Znf_Box.
DR InterPro: IPR001841; Znf_finger.
DR Pfam: PF00025; arf; 1.
DR Pfam: PF00643; zf-Bbox; 1.
DR SMART: SM00177; ARF; 1.
DR SMART: SM00502; BRC; 1.
DR SMART: SM00336; BBOX; 2.
DR SMART: SM00184; RING; 1.
DR TIGRfams: TIGR00231; small_gfp; 1.
DR PROSITE: PS01019; ARF; 1.
DR PROSITE: PS50119; ZF_BBOX; 1.
DR PROSITE: PS00518; ZF_RING_1; 1.
DR PROSITE: PS50089; ZF_RING_2; 1.
KW GTP-binding; Zinc-finger.
FT ZN_FING 1 56
FT NON_TER 1
FT ZN_FING 1 56
FT ZN_FING 102 148
FT DOMAIN 370 554
FT NP_BIND 391 398
FT NP_BIND 434 438
FT NP_BIND 493 496
FT SEQUENCE 554 AA; 62187 MW; F8427D6F27680839 CRC64;
SQ
Query Match 51.8%; Score 43; DB 1; Length 554;
Best Local Similarity 28.6%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 2; Indels 16; Gaps 1;
OY 2 GVC-----CGYKLCXHC 13
DB 12 GVCEDVPSLOGKVPRLLCGHTVCHDC 39
RESULT 9
ARDL_HUMAN
ID ARDL_HUMAN STANDARD: PRT; 574 AA.
AC P36406; Q9BZY5; Q9BZY4;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GTP-binding protein Arp-1 (Tripartite motif protein 23).
GN ARPD1 OR ARD1 OR TRIM23.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX MEDLINE=93232038; PubMed=8473324;
RA Mishima K., Tsuchiyama M., Nightingale M.S., Moss J., Vaughan M.;
RT "ARD 1, a 64-kDa guanine nucleotide-binding protein with a carboxyl-
terminal ADP-ribosylation factor domain."
RL J. Biol. Chem. 268:8801-8807(1993).
-1- FUNCTION: NOT KNOWN. THE C-TERMINUS CAN ACT AS AN ALLOSTERIC
ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT.
CC -1- FUNCTION: NOT KNOWN. THE C-TERMINUS CAN ACT AS AN ALLOSTERIC
ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT.
CC -----
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC TISSUE=Brain;
RA Strussberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NOT KNOWN. THE C-TERMINUS CAN ACT AS AN ALLOSTERIC
ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT.
CC -----
```


RA Shpakovskii G.V., Usery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: MAY NEGATIVELY REGULATE THE BASAL AND ACTIVATED
 CC TRANSCRIPTION OF MANY GENES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -1- SIMILARITY: CONTAINS 1 C3H1-TYPE ZINC FINGER.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 CC -1- SIMILARITY: STRONG, TO YEAST NOT4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: Z54366; CAA91192.1; -
 CC InterPro: IPR000504; RNA_rec_mot.
 CC InterPro: IPR003954; RRM_1.
 CC InterPro: IPR000571; Znf_CCH.
 CC InterPro: IPR001841; Znf_Ring.
 CC Pfam: PF00076; rrm; 1.
 CC DR Pfam: PF00642; zf_CCH; 1.
 CC DR SMART: SM00184; RING_1.
 CC DR SMART: SM00361; RRM_1; 1.
 CC DR PROSITE: PS50102; RRM; 1.
 CC DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 CC DR PROSITE: PS00518; zf_RING_1; FALSE_NEG.
 CC DR PROSITE: PS50089; zf_RING_2; 1.
 CC KW Hypothetical protein; Nuclear protein; Transcription regulation;
 CC KW Repressor; Zinc-finger; RNA-binding; Coiled coil.
 CC FT ZN_FING 18 61 RING-TYPE.
 CC FT DOMAIN 76 109 COILED COIL (POTENTIAL).
 CC FT DOMAIN 116 202 RNA-BINDING (RRM).
 CC SQ SEQUENCE 489 AA; 54417 MW; 237C2EF4D44221F7 CRC64;
 CC
 CC Query Match 50.6%; Score 42; DB 1; Length 489;
 CC Best Local Similarity 55.6%; Pred. No. 32;
 CC Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
 DB 37 CGYKLCXHC 45

RESULT 12

NOT4_YEAST STANDARD: PRT; 587 AA.
 ID NOT4_YEAST P34909;
 DT 01-FEB-1994 (Rel. 28; Last sequence update)
 DT 01-FEB-1994 (Rel. 28; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE General negative regulator of transcription subunit 4.
 GN NOT4 OR MOT2 OR SSI1 OR SSI1 OR CCL1 OR YER068W.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetiales; Saccharomycetaceae; Saccharomyces.
 CC NCBI_TaxID=4932;
 CC
 CC (1)
 CC SEQUENCE FROM N.A.
 CC MEDLINE=94313986; PubMed=8039500;
 CC Leberer E., Dignard D., Harcus D., Whiteway M., Thomas D.Y.;
 CC "Molecular characterization of Ssi1, a Saccharomyces cerevisiae gene
 CC involved in negative regulation of G-protein-mediated signal
 CC transduction.";
 CC EMBL J. 13:3050-3064(1994).
 CC
 CC (2)
 CC SEQUENCE FROM N.A.
 CC MEDLINE=94217711; PubMed=8164669;
 CC Cade R.M., Eirde B.;

RT "MOT2 encodes a negative regulator of gene expression that affects
 RT basal expression of pheromone-responsive genes in Saccharomyces
 RT cerevisiae.";
 RL Mol. Cell. Biol. 14:3139-3149(1994).
 CC [3]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=94217712; PubMed=8164670;
 CC Irie K., Yamaguchi K., Kawase K., Matsumoto K.;
 CC "The yeast MOT2 gene encodes a putative zinc finger protein that
 CC serves as a global negative regulator affecting expression of several
 CC categories of genes, including mating-pheromone-responsive genes.";
 CC Mol. Cell. Biol. 14:3150-3157(1994).
 CC [4]
 CC SEQUENCE FROM N.A.
 CC STRAIN=S288C / AB972;
 CC Dierrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 CC Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 CC Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
 CC Hyman R., Kayser A., Komp C., Lasikari D., Lew H., Lin D.,
 CC Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 CC Petel F.X., Roberts D., Sehl P., Schramm S., Shogen T., Smith V.,
 CC Taylor P., Wei Y., Yellon M., Bolstein D., Davis R.W.;
 CC Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
 CC [5]
 CC CHARACTERIZATION.
 CC MEDLINE=95011559; PubMed=7926748;
 CC Collart M.A., Struhl K.;
 CC "Nott1(CDC39), Not2(CDC36), Not3, and Not4 encode a global-negative
 CC regulator of transcription that differentially affects TATA-element
 CC utilization.";
 CC Genes Dev. 8:525-537(1994).
 CC -1- FUNCTION: NEGATIVELY REGULATES THE BASAL AND ACTIVATED
 CC TRANSCRIPTION OF MANY GENES. PREFERENTIALLY AFFECTS TC-TYPE TATA
 CC ELEMENT-DEPENDENT TRANSCRIPTION. COULD DIRECTLY OR INDIRECTLY
 CC INHIBITS COMPONENT(S) OF THE GENERAL TRANSCRIPTION MACHINERY.
 CC -1- SUBUNIT: FORMS A COMPLEX THAT COMPRISE NOT1, NOT2, NOT3, NOT4 AND
 CC NOT5.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 CC -1- SIMILARITY: STRONG, TO S.POMBE SPAC16C9.04C.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M96736; AAC37413.1; -
 CC EMBL: L26309; AAB00326.1; -
 CC EMBL: U18813; AAB64604.1; -
 CC PIR: S47918; S47918.
 CC PIR: S46366; S46366.
 CC TRANSFAC: T03591; -
 CC SGD: S0000870; MOT2.
 CC InterPro: IPR000504; RNA_rec_mot.
 CC InterPro: IPR003954; RRM_1.
 CC InterPro: IPR001841; Znf_Ring.
 CC Pfam: PF00076; rrm; 1.
 CC DR SMART: SM00184; RING_1.
 CC DR SMART: SM00361; RRM_1; 1.
 CC DR PROSITE: PS50102; RRM; 1.
 CC DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 CC DR PROSITE: PS00518; zf_RING_1; FALSE_NEG.
 CC DR PROSITE: PS50089; zf_RING_2; 1.
 CC KW Nuclear protein; Transcription regulation; Repressor; Zinc-finger;
 CC RNA-binding; Coiled coil.
 CC FT ZN_FING 33 78 RING-TYPE.
 CC FT DOMAIN 94 128 COILED COIL (POTENTIAL).
 CC FT DOMAIN 137 232 RNA-BINDING (RRM).
 CC SQ SEQUENCE 587 AA; 65354 MW; 8847084BD9BF48B7 CRC64;

SQ SEQUENCE 172 AA; 19418 MW; 9DA1C817CA36C8B9 CRC64;

Query Match 49.4%; Score 41; DB 1; Length 172;
Best Local Similarity 60.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 GVC CGKLC 11
1 1 1 1 1 1
DB 67 GICIGVLC 76

RESULT 15
Y653_METJA

ID Y653_METJA STANDARD; PRT; 194 AA.

AC 058069;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0653.
GN MJ0653.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.V., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).

CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U67513; AAB98648.1; -.
DR TIGR; MJ0653; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 2.
DR SMART; SM00116; CBS; 2.
KW Hypothetical protein; Repeat; CBS domain; Complete proteome.
FT DOMAIN 9 63 CBS 1.
FT DOMAIN 76 127 CBS 2.
FT DOMAIN 187 194 GCU-RICH
SQ SEQUENCE 194 AA; 21723 MW; 33921320EF82E5DC CRC64;

Query Match 49.4%; Score 41; DB 1; Length 194;
Best Local Similarity 36.0%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 2; Indels 12; Gaps 2;

OY 1 NGVC--CGTK-----LCHXC 13
1 1 1 1 1 1 1 1
DB 161 NGICENCGYGRVRLXQGRVLCDEC 185

Search completed: December 2, 2002, 10:07:27
Job time : 8.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:05:01 ; Search time 24.5 seconds
(without alignments)
109.331 Million cell updates/sec

Title: US-09-787-986a-1
Perfect score: 83
Sequence: 1 NGVCCGKXCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	68.7	258	12	036453 lymantria d
2	56	67.5	151	12	010614 hellicoverpa
3	56	67.5	285	12	099h38 hellicoverpa
4	56	67.5	285	12	08v5x4 hellicoverpa
5	54	65.1	244	12	091B09 spodoptera
6	51	61.4	234	12	080166 manestera co
7	49	59.0	289	12	091BK7 spodoptera
8	48	57.8	64	5	09BPE9 conus penna
9	48	57.8	514	5	095QT3 caenorhabd
10	47	56.6	112	12	081680 hepatitis c
11	47	56.6	322	10	039896 glycine max
12	47	56.6	326	10	039895 glycine max
13	47	56.6	364	12	081547 hepatitis c
14	47	56.6	373	10	09M1P9 arabidopsis
15	46	55.4	69	2	09R6Z5 nocardioid
16	46	55.4	121	11	09CZW7 mus musculu

17	46	55.4	372	11	099ND4	099nd4 ratus norv
18	46	55.4	538	4	096IF1	096if1 homo sapien
19	46	55.4	547	11	P97472	P97472 mus musculu
20	46	55.4	547	11	091XC0	091xc0 mus musculu
21	45	54.2	79	16	097P38	097p38 streptococ
22	45	54.2	220	5	08S059	08s059 encephalito
23	45	54.2	326	16	08RA96	08ra96 thermoaer
24	45	54.2	446	4	096B03	096bq3 homo sapien
25	45	54.2	721	12	091LL5	091ll5 white spot
26	45	54.2	1009	12	08VAC3	08vac3 white spot
27	44	53.0	117	17	08TJD7	08tjd7 methanosarc
28	44	53.0	144	17	08TR68	08tr68 methanosarc
29	44	53.0	144	17	08TPM7	08tpm7 methanosarc
30	44	53.0	144	17	08TNM6	08tnm6 methanosarc
31	44	53.0	144	17	08TNM3	08tnm3 methanosarc
32	44	53.0	144	17	08TLT2	08ltl2 methanosarc
33	44	53.0	144	17	08TKT4	08kt4 methanosarc
34	44	53.0	144	17	08TH59	08th59 methanosarc
35	44	53.0	144	17	08TH58	08th58 methanosarc
36	44	53.0	144	17	08TH57	08th57 methanosarc
37	44	53.0	190	17	08TMC6	08tmc6 methanosarc
38	44	53.0	243	12	091GD2	091gd2 epiptyas po
39	44	53.0	244	5	062463	062463 caenorhabd
40	44	53.0	244	5	062464	062464 caenorhabd
41	44	53.0	332	5	018012	018012 caenorhabd
42	44	53.0	465	4	060260	060260 homo sapien
43	44	53.0	514	10	09LD86	09ld86 arabidopsis
44	44	53.0	600	11	060813	060813 mus musculu
45	44	53.0	789	11	P70505	P70505 ratus norv

ALIGNMENTS

RESULT 1

036453 PRELIMINARY: PRT; 258 AA.

AC 036453: 09YMW3:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Immediate early 0 protein (Immediate early transactivator 0).
IE-0.
OS Lymantria dispar multicapsid nuclear polyhedrosis virus (LdMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.
OX Nucleopolydnavirus.
OC NCBI_TaxId=10449;
RX MEDLINE-97445058: PubMed-9300047;
RN SEQUENCE FROM N.A.
RT Pearson M.N., Rohmann G.F.;
RT Splicing is required for transactivation by the immediate early gene
RT 1 of the Lymantria dispar multinucleocapsid nuclear polyhedrosis
RT virus.";
RL Virology 235:153-165(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-9124785: PubMed-9887315;
RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA Slavacek J.M., Rohmann G.F.;
RT *Sequence and analysis of the genome of a baculovirus pathogenic for
RT Lymantria dispar".
RL Virology 253:17-34(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA Slavacek J., Rohmann G.F.;
RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX Kuzio J.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF006656; AAC58234.1; -.

DR EMBL; AF081810; AAC70206.1; -.
 DR InterPro; IPR001841; znf_ring.
 DR SMART; SM00184; RING; 1.
 SQ SEQUENCE 258 AA; 29395 MW; ACS64CDF928282BAD CRC64;

Query Match 68.7%; Score 57; DB 12; Length 258;
 Best Local Similarity 63.6%; Pred. No. 0.069;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 CCGYKLCCHXC 13
 |||||:|:|
 Db 208 CCGYKLCCHXC 218

RESULT 2

ID 010614 PRELIMINARY; PRT; 151 AA.
 AC 010614;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE AcMNPV ORF1 homolog.
 OS Helicoverpa zea single nucleocapsid nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10468;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ELKAR;
 RX MEDLINE-97321796; PubMed-9178498;
 RA Le T.H., Wu T., Robertson A., Bulach D., Cowan P., Goodge K.,
 RA Trilhe D.;
 RT "Genetically variable triplet repeats in a RING-finger ORF of
 RT Helicoverpa species baculoviruses.";
 RL Virus Res 49:67-77(1997).
 DR EMBL; U67264; AAS54095.1; -.
 DR InterPro; IPR001841; znf_ring.
 DR SMART; SM00184; RING; 1.
 SQ SEQUENCE 151 AA; 17548 MW; 4986432F6DCD3169 CRC64;

Query Match 67.5%; Score 56; DB 12; Length 151;
 Best Local Similarity 70.0%; Pred. No. 0.064;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 CCGYKLCCHXC 13
 |||||:|:|
 Db 101 CCGYKLCCHXC 110

RESULT 3

ID 099H38 PRELIMINARY; PRT; 285 AA.
 AC 099H38;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Ie-0.
 OS Helicoverpa armigera nucleopolyhedrovirus G4, and
 OS Helicoverpa armigera nuclear polyhedrosis virus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=148363; 51313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
 RA Deng F., Chen X., Vlak J.M., Arif B.M., Hu Z.;
 RT "Sequence analysis of the gp37 gene of Heliothis armigera single-
 RT nucleocapsid nucleopolyhedrovirus.";
 RT Zhongguo Bingduxue 15:35-42(2000).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
 RA Wang H., Hu Z., Sun X., Vlak J.M., Chen X.;

RT "Sequence analysis of the iap3 gene of Heliothis armigera single-
 RT nucleocapsid nucleopolyhedrovirus.";
 RT Zhongguo Bingduxue 15:43-49(2000).
 RL [3]

RP SEQUENCE FROM N.A.
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
 RX MEDLINE-21078302; PubMed-11210934;
 RA Wang H., Chen X., Wang H., Arif B.M., Vlak J.M., Hu Z.;
 RA "Nucleotide sequence and transcriptional analysis of a putative basic
 RT DNA-binding protein of Helicoverpa armigera polyhedrovirus.";
 RL Virus Genes 22:113-120(2001).

[4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
 RX MEDLINE-21064569; PubMed-11125177;
 RA Chen X., Irtkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
 RA Peters S., Zuidema D., Lankhorst R.K., Vlak J.M., Hu Z.;
 RT "The sequence of the Helicoverpa armigera single-nucleocapsid
 RT nucleopolyhedrovirus genome.";
 RL J. Gen. Virol. 82:241-257(2001).

[5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
 RA Chen X., Irtkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
 RA Peters S., Zuidema D., Lankhorst R.K., Vlak J.M., Hu Z.;
 RL Submitted (May-2000) to the EMBL/Genbank/DBJ databases.

[6]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Helicoverpa armigera nuclear polyhedrosis virus; STRAIN-C1;
 RA Zhang C.X., Wu J.C.;
 RT "Genome structure and the p10 gene of the Helicoverpa armigera
 RT nucleopolyhedrovirus.";
 RN Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 33:179-184(2001).

[7]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Helicoverpa armigera nuclear polyhedrosis virus; STRAIN-C1;
 RA Zhang C.X., Jin W.R.;
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF271059; AAG53751.1; -.
 DR EMBL; AF303045; AAK96301.1; -.
 DR InterPro; IPR001841; znf_ring.
 DR SMART; SM00184; RING; 1.

SQ SEQUENCE 285 AA; 33186 MW; C5FC3AE65BA27BDD CRC64;

Query Match 67.5%; Score 56; DB 12; Length 285;
 Best Local Similarity 70.0%; Pred. No. 0.11;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 CCGYKLCCHXC 13
 |||||:|:|
 Db 235 CCGYKLCCHXC 244

RESULT 4

ID 08V5X4 PRELIMINARY; PRT; 285 AA.
 AC 08V5X4;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE ORF8
 OS Helicoverpa zea single nucleocapsid nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10468;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chen X., Zhang W.-J., Wong J., Chun G., Lu A., McCutchen B.F.,
 RA Presnail J.K., Herimann R., Dolan M., Tingey S., Hu Z.-H., Vlak J.M.;
 RT "Genome sequence analysis of Helicoverpa zea single nucleocapsid
 RT nucleopolyhedrovirus.";
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF334030; AAL56153.1; -.

DR InterPro: IPR001841; Znf_ring.
DR PROSITE: PS50089; ZF_RING.2; 1.
SQ SEQUENCE 285 AA; 33189 MW; EF9E35A71B8E7F3D CRC64;

Query Match 67.5%; Score 56; DB 12; Length 285;
Best Local Similarity 70.0%; Pred. No. 0.11;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
DB 235 CCGYKLCXHC 244

RESULT 5

OY 091B09 PRELIMINARY; PRT; 244 AA.
AC 091B09;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ORF138 Ie0.
OS Spodoptera exigua nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10454;
[1]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=97437494; PubMed=9292027;
RA van Strien E.A., Faktor O., Hu Z.H., Zuidema D., Goldbach R.W.,
RA Vlak J.M.;
RT "Baculoviruses contain a gene for the large subunit of ribonucleotide
RT reductase."
RL J. Gen. Virol. 78:2365-2377(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20036646; PubMed=10567663;
RA IJkel W.F., van Strien E.A., Heldens J.G., Broer R., Zuidema D.,
RA Goldbach R.W., Vlak J.M.;
RT "Sequence and organization of the spodoptera exigua multicapsid
RT nucleopolyhedrovirus genome."
RL J. Gen. Virol. 80:3289-3304(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Zuidema D.;
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA IJkel W.F., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D.,
RA Goldbach R.W., Vlak J.M.;
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF169823; AAF33667.1; -;
DR InterPro: IPR001841; Znf_ring.
DR SMART: SM00184; RING; 1.
SQ SEQUENCE 244 AA; 28719 MW; 1F7662E837A866DB CRC64;

Query Match 65.1%; Score 54; DB 12; Length 244;
Best Local Similarity 70.0%; Pred. No. 0.21;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
DB 199 CCGYKLCXHC 208

RESULT 6

OY 080L66 PRELIMINARY; PRT; 234 AA.
AC 080L66;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE IE0.
OS Mamestra configurata nucleopolyhedrovirus.

OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=191492;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=90/2;
RX MEDLINE=97163493; PubMed=9010313;
RA Li S., Erlandson M., Moody D., Gillott C.;
RT "A physical map of the Mamestra configurata nucleopolyhedrovirus
RT genome and sequence analysis of the polynedrin gene."
RL J. Gen. Virol. 78:265-271(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=90/2;
RX MEDLINE=21884635; PubMed=11886270;
RA Li Q., Donly C., Li L., Willis L.G., Theilmann D.A., Erlandson M.;
RT "Sequence and Organization of the Mamestra configurata
RT Nucleopolyhedrovirus Genome."
RL Virology 294:106-121(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=90/2;
RA Li Q., Donly C., Li L., Willis L.G., Theilmann D.A., Erlandson M.A.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: U59461; AAM09276.1; -;
SQ SEQUENCE 234 AA; 27232 MW; 5E5F1330CD7A711 CRC64;

Query Match 61.4%; Score 51; DB 12; Length 234;
Best Local Similarity 60.0%; Pred. No. 0.63;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
DB 193 CCGYKLCXHC 202

RESULT 7

OY 091BK7 PRELIMINARY; PRT; 289 AA.
AC 091BK7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Immediate early 0 protein.
OS Spodoptera litura nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46242;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=G2;
RX MEDLINE=21425398; PubMed=11531416;
RA Pang Y., Yu J., Wang L., Hu X., Bao W., Li G., Chen C., Han H., Hu S.,
RA Yang H.;
RT "Sequence Analysis of the Spodoptera litura Multicapsid
RT Nucleopolyhedrovirus Genome."
RL Virology 287:391-404(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=G2;
RA Yu J., Wang L., Hu X., Pang Y.;
RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: AF325155; AAL01695.1; -;
DR InterPro: IPR001841; Znf_ring.
DR Zinc-finger.
KW ZINC-FINGER.
SQ SEQUENCE 289 AA; 33387 MW; 50C3F3E63FE78C6B CRC64;

Query Match 59.0%; Score 49; DB 12; Length 289;
Best Local Similarity 50.0%; Pred. No. 1.6;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13

Db 245 CCGFRICNLIC 254

RESULT 8

Q9BPE9 PRELIMINARY; PRT; 64 AA.
ID Q9BPE9;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
OS Conotoxin scafold 1X.
OC Conus pennaeus (feathered cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conidae; Conus.
OX NCBI_Taxid=37335;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conicello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Falznalber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides."
RL MOL. BIOL. EVOL. 18:120-131(2001).
DR EMBL; AF214980; AAG60408.1; -;
SQ SEQUENCE 64 AA; 6928 MW; 0AB87620FCCC1410 CRC64;

Query Match 57.8%; Score 48; DB 5; Length 64;
Best Local Similarity 60.0%; Pred. No. 0.65;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CCGYKLCXHC 13
ID 54 CCGYKLCXHC 63

RESULT 9

Q95QT3 PRELIMINARY; PRT; 514 AA.
ID Q95QT3;
AC Q95QT3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 58.6 kDa protein.
GN C28G1.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;

"Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
Science 282:2012-2018(1998).
[2]

SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Favello A.;
RT "The sequence of C. elegans cosmid C28G1.";
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
RN [3]

SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 RING-TYPE ZINC FINGERS.
DR EMBL; U41026; AAL02448.1; -;
DR InterPro; IPR000315; Znf_Box.
DR InterPro; IPR001841; Znf_Zinc.
DR Pfam; PF00643; zf-B_box; 1.

DR Pfam; PF00097; zf-C3HC4; 2.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_2.
KW Hypothetical protein; Zinc-finger.
SQ SEQUENCE 514 AA; 58600 MW; BC33388F0F599446 CRC64;

Query Match 57.8%; Score 48; DB 5; Length 514;
Best Local Similarity 63.6%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 VCCGYKLCXHC 13
ID 263 VTCGHALCHXC 273

RESULT 10

Q81680 PRELIMINARY; PRT; 112 AA.
ID Q81680;
AC Q81680;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NE048;
RA Tokita H., Okamoto H.;
RL Submitted (JAN-1993) to the EMBL/Genbank/DBJ databases.
DR EMBL; D14198; BAA03233.1; -;
DR InterPro; IPR002166; HCV_RdRP.
DR Pfam; PR00998; HCV_RdRP; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 112 112
FT 112 112
SQ SEQUENCE 112 AA; 12008 MW; D7D79CA732ED9D3F CRC64;

Query Match 56.6%; Score 47; DB 12; Length 112;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVCQGYKLC 10
ID 45 GLCCGYRRC 53

RESULT 11

Q39896 PRELIMINARY; PRT; 322 AA.
ID Q39896;
AC Q39896;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE TGAGC-motif-binding factor.
GN STF2.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_Taxid=3847;
RN [1]

SEQUENCE FROM N.A.
RP STRAIN-WILLIAMS; TISSUE=HYPOCOTYL;
RX MEDLINE=98388650; PubMed=9721678;
RA Cheong Y.H., Yoo C.M., Park J.M., Ryu G.R., Goekjian V.H., Nagao R.T.,
RA Key J.L., Cho M.J., Hong J.C.;
RT "STF1 is a novel TGAGC-binding factor with a zinc-finger motif and a
bZIP domain which heterodimerizes with GSF proteins.";
RL Plant J. 15:199-209(1998).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE bZIP FAMILY.

DR EMBL: L28004; AAC05018.1; -
DR TRANSFAC: T02973; -
DR InterPro: IPR004827; TF_bZIP.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF00170; bZIP; 1.
DR SMART: SM00338; BRLZ; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00036; bZIP_BASIC; 1.
DR DNA-binding: Nuclear protein.
SQ SEQUENCE 322 AA; 35238 MW; BE692E01EF6FA6E5 CRC64;

Query Match 56.6%; Score 47; DB 10; Length 322;
Best Local Similarity 66.7%; Pred. No. 3.8;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
DB 61 CGFPLCHSC 69

RESULT 12
039895 PRELIMINARY; PRT; 326 AA.
ID 039895
AC 039895
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE TGACG-motif binding factor.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WILLIAMS; TISSUE-HYPOCOTYL;
RX MEDLINE=98388650; PubMed=9721678;
RA Cheong Y.H., Yoo C.M., Park J.M., Ryu G.R., Goekjhan V.H., Nagao R.T.,
RA Key J.L., Cho M.J., Hong J.C.;
RT *STF1 is a novel TGACG-binding factor with a zinc-finger motif and a
RT bZIP domain which heterodimerizes with GBF proteins.*;
RL Plant J. 15:199-209(1998).
CC -1 SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1 SIMILARITY: BELONGS TO THE bZIP FAMILY.
DR EMBL: L28003; AAC05017.1; -
DR TRANSFAC: T02969; -
DR InterPro: IPR004827; TF_bZIP.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF00170; bZIP; 1.
DR SMART: SM00338; BRLZ; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00036; bZIP_BASIC; 1.
DR DNA-binding: Nuclear protein.
KW SEQUENCE 326 AA; 35488 MW; 0BF960434DC19AC5 CRC64;

Query Match 56.6%; Score 47; DB 10; Length 326;
Best Local Similarity 66.7%; Pred. No. 3.8;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
DB 61 CGFPLCHSC 69

RESULT 13
081547 PRELIMINARY; PRT; 364 AA.
ID 081547
AC 081547
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NE048;
RX MEDLINE=94201770; PubMed=8151307;
RA Tokita H., Shrestha S.M., Okamoto H., Sakamoto M., Horikita M.,
RA Iizuka H., Shrestha S., Miyakawa Y., Mayumi M.,
RT "Hepatitis C virus variants from Nepal with novel genotypes and their
RT classification into the third major group.*";
RL J. Gen. Virol. 75:931-936(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NE048;
RA Okamoto H.;
RL Submitted (MAY-1994) to the EMBL/Genbank/DBJ databases.
DR EMBL: D16613; BAA04035.1; -
DR InterPro: IPR002166; HCV_RDRP.
DR Pfam: PF00998; HCV_RDRP; 1.
FT Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
SQ SEQUENCE 364 AA; 39979 MW; 9EE13125B350EF12 CRC64;

Query Match 56.6%; Score 47; DB 12; Length 364;
Best Local Similarity 66.7%; Pred. No. 4.2;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVC CGYKLC 10
DB 44 GLCCGYRRC 52

RESULT 14
09MIF9 PRELIMINARY; PRT; 373 AA.
ID 09MIF9
AC 09MIF9
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 42.7 kDa protein.
GN F9K21.60.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Jordan N., Bangert S., Wiedemann R., Voss H., Unseld M., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetler F., Salanoubat M.;
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL138657; CAB75477.1; -
DR InterPro: IPR002106; AATRNA_LigaseII.
DR InterPro: IPR002867; Znf_C6HC.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF01485; IIR; 2.
DR SMART: SM00184; RING; 2.
DR PROSITE: PS00179; AA_TRNA_LIGASE_II_1; UNKNOWN_1.
DR PROSITE: PS00518; ZF_RING_1; 1.
KW Hypothetical protein; zinc-finger.
SQ SEQUENCE 373 AA; 42688 MW; A4484B5BF0667BB CRC64;

Query Match 56.6%; Score 47; DB 10; Length 373;
Best Local Similarity 66.7%; Pred. No. 4.3;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
DB 61 CGFPLCHSC 69

Db 329 CGKRCYAC 337

RESULT 15

```
Q9R6Z5 ID Q9R6Z5 PRELIMINARY; PRT; 69 AA.
AC Q9R6Z5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FERREDOXIN.
GN PHDC.
OS Nocardioides sp. (strain KP7), and
OC Nocardioides sp. KP7.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Propionibacterineae; Nocardioidaceae; Nocardioides.
OX NCBI_TaxID=35761, 102632;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Nocardioides sp.; STRAIN=KP7;
RX MEDLINE=99169927; PubMed=10070721;
RA Saito A., Iwabuchi T., Harayama S.;
RT "Characterization of genes for enzymes involved in the phenanthrene
  degradation in Nocardioides sp. KP7."
RL Chemosphere 38:1331-1337(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Nocardioides sp. KP7; STRAIN=KP7;
RX MEDLINE=20200350; PubMed=10735855;
RA Saito A., Iwabuchi T., Harayama S.;
RT "A novel phenanthrene dioxygenase from Nocardioides sp. strain KP7:
  Expression in escherichia coli."
RL J. Bacteriol. 182:2134-2141(2000).
DR EMBL: AB017795; BA84714.1; -
DR EMBL: AB031319; BA94713.1; -
DR HSSP: P46797; IVW.
DR InterPro: IPR001080; 3Fe4S_ferredoxin.
DR PRINTS: PR00352; 3FE4SFRDOXIN.
SQ SEQUENCE 69 AA; 7348 MW; 288EA465E055E3E1 CRC64;
```

Query Match 55.4%; Score 46; DB 2; Length 69;
Best Local Similarity 85.7%; Pred. No. 1.5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 CCGKRLC 10
| | | | |
DB 10 CCGKRLC 16

Search completed: December 2, 2002, 10:08:24
Job time : 25.5 secs

PI Lewis RJ, Alewood PF, Sharpe IA, ...

Cysteine-added var
Novel human diagno
Protophila melango
Benzene ring hydr
Novel human polype
Human Ajba-like p
Novel human diagno
Novel human diagno
Human haemochrom
Human NF- κ B activat
Human NF- κ B activ
Shrimp white spot
Human OREX protein
Human SSA-56kDa p
Herbically active
Human gastric cano
Human secreted pro
Granulopietic act
Human G-CSF mutat
Human G-CSF mutat
Rat AD 1. Ratlus
Human ARD 1. Homo
Streptococcus poly
Human protein SEO
Novel human diagno
Novel human diagno
Novel human diagno
Novel human diagno
Novel human diagno
Novel human diagno
Novel human diagno
Omega-conopeptide
Omega-conopeptide
Omega-conopeptide

DR WPI: 2000-303738/26.
XX
PT Isolated, synthetic or recombinant chi-conotoxin peptide capable of
PT inhibiting neuronal amine transporter used for treatment or prophylaxis
PT of urinary or cardiovascular conditions, mood disorders, or
PT treatment/control of pain/inflammation
PS
XX Claim 3; Page 33; 47pp; English.
XX
CC This conotoxin, chi-Mr1B, is a member of a new class of conotoxins,
CC designated chi-conotoxin. It was isolated from the venom of the mollusc
CC hunting cone snail, Conus marmoreus. The peptide is an inhibitor of the
CC neuronal amine transporters, especially the neuronal norepinephrine
CC transporter. Inhibitors of norepinephrine re-uptake which have a
CC negligible anti-cholinergic effect are particularly useful in the
CC treatment of lower urinary tract disorders. Chi-Mr1A (0.1 nM-1 micro M)
CC inhibited the accumulation of radiolabeled norepinephrine in a
CC concentration-dependent manner, with a log IC-50 value of -8.17 plus or
CC minus 0.0275 (n = 4). The concentration of chi-Mr1A required to inhibit
CC the accumulation by 50 percent was found to be approximately 7 nM. This
CC concentration is approximately one order of magnitude lower than that
CC needed for desipramine to produce the same effect. The peptides are
CC useful for the treatment or prophylaxis of urinary or cardiovascular
CC conditions or diseases (arrhythmia or coronary heart failure) or mood
CC disorders (depression, anxiety or cravings), or the treatment or control
CC of pain or inflammation (chronic pain, neuropathic pain or inflammatory
CC pain).
XX
SQ Sequence 13 AA:
Query Match 96.3%; Score 78; DB 21; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.00058;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GVCCGKRLCHXC 13
DB 1 GVCCGKRLCHPC 13
RESULT 2
AAB08018
ID AAB08018 standard; peptide; 12 AA.
XX
AC AAB08018;
XX
DT 14-NOV-2000 (first entry)
XX
DE Amino acid sequence of the conotoxin peptide Mar2.
XX
KM Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Mar2.
XX
OS Conus marmoreus.
XX
FH Key Location/Qualifiers
FT Modified-site 12 /note="hydroxy-Pro"
XX
PN W0200044769-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000MO-US01978.
XX
XX 29-JAN-1999; 99US-0118381.
PR 28-DEC-1999; 99US-0173343.
XX
PA (UTAH) UNIV UTAH RES FOUND.
XX
PI McIntosh JM, Olivera BM, Cruz LJ;
XX WPI: 2000-476222/41.
DR
PT Purified₂ap-conotoxin derived from cone snail venom for use as an

PT analgesic -
XX
XX Claim 14; Page -: 29pp; English.
XX
CC The present sequence represents an ap-conotoxin peptide, designated
CC Mar2. Conotoxins are naturally available in minute amounts in the
CC venom of cone snails. The peptides have analgesic activity. The
CC peptides are used to treat or prevent pain.
CC note: this sequence does not appear in the specification; it was created
CC using information provided.
XX
SQ Sequence 12 AA:
Query Match 91.4%; Score 74; DB 21; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0018;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 GVCCGKRLCHXC 13
DB 1 GVCCGKRLCHPC 12
RESULT 3
AAB08017
ID AAB08017 standard; peptide; 13 AA.
XX
AC AAB08017;
XX
DT 14-NOV-2000 (first entry)
XX
DE Amino acid sequence of the conotoxin peptide Mar1.
XX
KM Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Mar1.
XX
OS Conus marmoreus.
XX
FH Key Location/Qualifiers
FT Modified-site 12 /note="hydroxy-Pro"
XX
PN W0200044769-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000MO-US01978.
XX
XX 29-JAN-1999; 99US-0118381.
PR 28-DEC-1999; 99US-0173343.
XX
PA (UTAH) UNIV UTAH RES FOUND.
XX
PI McIntosh JM, Olivera BM, Cruz LJ;
XX WPI: 2000-476222/41.
DR
PT Purified ap-conotoxin derived from cone snail venom for use as an
PT analgesic -
XX
PS Claim 13; Page -: 29pp; English.
XX
CC The present sequence represents an ap-conotoxin peptide, designated
CC Mar1. Conotoxins are naturally available in minute amounts in the
CC venom of cone snails. The peptides have analgesic activity. The
CC peptides are used to treat or prevent pain.
CC note: this sequence does not appear in the specification; it was created
CC using information provided.
XX
SQ Sequence 13 AA:
Query Match 91.4%; Score 74; DB 21; Length 13;
Best Local Similarity 91.7%; Pred. No. 0.0019;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVCCGYKLCXHC 13
 |||||
 DB 2 GVCCGYKLCXHC 13

RESULT 4
 AAY92229
 ID AAY92229 standard; peptide; 13 AA.
 XX AAY92229;

10-AUG-2000 (first entry)

Chi-conotoxin peptide, chi-MrIA.

chi-conotoxin; chi-MrIA; cone snail; inhibitor; amine transporter;
 neuronal; noradrenaline transporter; urinary tract disorder; analgesic;
 antiarrhythmic; cardiac; antidepressant; anxiolytic; anti-inflammatory.
 Conus marmoreus.

Key Location/Qualifiers
 Misc-difference 12
 /label= 4HYP
 /note="4-hydroxyproline"

MO200020444-A1.

13-APR-2000.

01-OCT-1999; 99WO-AU00844.

02-OCT-1998; 98AU-0006274.

(UYOU) UNIV QUEENSLAND.

Lewis RJ, Alewood PF, Sharpe JA;

WPI: 2000-303738/26.

Isolated, synthetic or recombinant chi-conotoxin peptide capable of
 inhibiting neuronal amine transporter used for treatment or prophylaxis
 of urinary or cardiovascular conditions, mood disorders, or
 treatment/control of pain/inflammation

Claim 3; Page 33; 47pp; English.

This conotoxin, chi-MrIA, is a member of a new class of conotoxins,
 designated chi-conotoxin. It was isolated from the venom of the mollusc
 hunting cone snail, Conus marmoreus. The peptide is an inhibitor of the
 neuronal amine transporters, especially the neuronal noradrenaline
 transporter. Inhibitors of noradrenaline re-uptake which have a
 negligible anti-cholinergic effect are particularly useful in the
 treatment of lower urinary tract disorders. Chi-MrIA (0.1 nm-1 micro M)
 inhibited the accumulation of radiolabeled noradrenaline in a
 concentration-dependent manner, with a log IC-50 value of -8.17 plus or
 minus 0.0275 (n = 4). The concentration of chi-MrIA required to inhibit
 the accumulation by 50 percent was found to be approximately 7 nM. This
 concentration is approximately one order of magnitude lower than that
 needed for desipramine to produce the same effect. The peptides are
 useful for the treatment or prophylaxis of urinary or cardiovascular
 conditions or diseases (arrhythmia or coronary heart failure) or mood
 disorders (depression, anxiety or cravings), or the treatment or control
 of pain or inflammation (chronic pain, neuropathic pain or inflammatory
 pain).

Sequence 13 AA;

Query Match 91.4%; Score 74; DB 21; Length 13;

Best Local Similarity 91.7%; Pred. No. 0.0019; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVCCGYKLCXHC 13

DB 2 GVCCGYKLCXHC 13
 |||||

RESULT 5
 AAB08016
 ID AAB08016 standard; Protein; 61 AA.
 XX AAB08016;

14-NOV-2000 (first entry)

Amino acid sequence of the conotoxin Marl propeptide.

Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.

Conus marmoreus.

MO200044769-A1.

03-AUG-2000.

28-JAN-2000; 2000WO-US01978.

29-JAN-1999; 99US-0118381.

28-DEC-1999; 99US-0173343.

(UTAH) UNIV UTAH RES FOUND.

McIntosh JM, Olivera BM, Cruz LJ;

WPI: 2000-476222/41.

N-PSDB: AAA63513.

Purified ap-conotoxin derived from cone snail venom for use as an
 analgesic -

Claim 23; Page 13-14; 29pp; English.

The present sequence represents a Marl propeptide. Marl is an
 ap-conotoxin peptide. Conotoxins are naturally available in minute
 amounts in the venom of cone snails. The peptides have analgesic
 activity. The peptides are used to treat or prevent pain.

Sequence 61 AA;

Query Match 91.4%; Score 74; DB 21; Length 61;
 Best Local Similarity 91.7%; Pred. No. 0.0067; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVCCGYKLCXHC 13
 |||||
 DB 50 GVCCGYKLCXHC 61

RESULT 6
 AAY92231

ID AAY92231 standard; Protein; 61 AA.

AC AAY92231;

10-AUG-2000 (first entry)

Chi-conotoxin, chi-MrIA, leader and mature peptide.

chi-conotoxin; chi-MrIA; cone snail; inhibitor; amine transporter;

neuronal; noradrenaline transporter; urinary tract; analgesic; cardiac;
 antiarrhythmic; antidepressant; anxiolytic; anti-inflammatory.

Conus marmoreus.

MO200020444-A1.

XX

XX	13-APR-2000.
PF	01-OCT-1999; 99WO-AU00844.
XX	
PR	02-OCT-1998; 98AU-0006274.
XX	
PA	(UYQU) UNIV QUEENSLAND.
XX	
PI	Lewis RJ, Alewood PF, Sharpe IA:
XX	
DR	WPI; 2000-303738/26.
DR	N-PsDB; AAA09112.
PT	Isolated, synthetic or recombinant chi-conotoxin peptide capable of inhibiting neuronal amine transporter used for treatment or prophylaxis of urinary or cardiovascular conditions, mood disorders, or treatment/control of pain/inflammation
XX	
PS	Example 7; Page 31; 47pp; English.
XX	
CC	This sequence is the conotoxin, chi-MRIA, a member of a new class of conotoxins, designated chi-conotoxin. It was isolated from the venom of the mollusc hunting cone snail, Conus marmoreus. The peptide is an inhibitor of the neuronal amine transporters, especially the neuronal noradrenaline transporter. Inhibitors of noradrenaline re-uptake which have a negligible anti-cholinergic effect are particularly useful in the treatment of lower urinary tract disorders. Chi-MRIA (0.1 nm-1 micro M) inhibited the accumulation of radiolabeled noradrenaline in a concentration-dependent manner, with a log IC-50 value of -8.17 plus or minus 0.0275 (n = 4). The concentration of chi-MRIA required to inhibit the accumulation by 50 percent was found to be approximately 7 nM. This concentration is approximately one order of magnitude lower than that needed for desipramine to produce the same effect. The peptides are useful for the treatment or prophylaxis of urinary or cardiovascular conditions or diseases (arrhythmia or coronary heart failure) or mood disorders (depression, anxiety or cravings), or the treatment or control of pain or inflammation (chronic pain, neuropathic pain or inflammatory pain).
CC	
XX	
SQ	Sequence 61 AA:
OY	Query Match 91.4%; Score 74; DB 21; Length 61; Best Local Similarity 91.7%; Pred. No. 0.0067;
Db	Matches 11: Conservative 0; Mismatches 1; Indels 0; Gaps 0
OY	2 GVCGGYKLCXHC 13
AB08019	AAB08019 standard; peptide; 12 AA.
AC	AAB08019;
DT	14-NOV-2000 (first entry)
DE	Amino acid sequence of the conotoxin peptide U036.
XX	
OS	Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; U036.
XX	
OS	Conus marmoreus.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 7 /label= Lys, Xaa
FT	/note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
FT	Modified-site 11 N,N,N-trimethyl Lys"
FT	/note= "hydroxy-Pro"
XX	
PN	WO200044769-A1.

```

XX 03-AUG-2000.
PD
XX
XX 28-JAN-2000; 2000OWO-US01978.
PF
XX 29-JAN-1999; 99US-0118381.
PR
XX 28-DEC-1999; 99US-0173343.
XX
PA (UTAH ) UNIV UTAH RES FOUND.
XX
XX McIntosh JM, Olivera BM, Cruz LJ;
PI
XX WPI: 2000-476222/41.
DR
XX
XX Purified ap-conotoxin derived from cone snail venom for use as an
PT analgesic -
XX
XX Claim 15; Page -: 29pp; English.
PS
XX
XX The present sequence represents an ap-conotoxin peptide, designated
CC U036. Conotoxins are naturally available in minute amounts in the
CC venom of cone snails. The peptides have analgesic activity. The
CC peptides are used to treat or prevent pain.
CC note: this sequence does not appear in the specification; it was created
CC using information provided.
CC
XX
SQ Sequence 12 AA;
Query Match 84.0%; Score 68; DB 21; Length 12;
Best Local Similarity 83.3%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

QY 2 GVC CGYKLCXHC 13
      ||||| |||
Db 1 GVCCGYXLCXHC 12

RESULT 8
AAB08014
ID AAB08014 standard; peptide; 12 AA.
XX
XX AAB08014;
AC
XX
XX 14-NOV-2000 (first entry)
DT
XX
XX Generic formula for conotoxin peptide Mar2.
DE
XX
XX Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Mar2.
KW
XX
XX Conus marmoreus.
OS
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /label= Tyr, Xaa
FT /note= "Xaa is mono-halo-Tyr, di-halo-Tyr"
FT O-sulfo-Tyr, O-phospho Tyr or nitro-Tyr"
FT Misc-difference 7 /label= Lys, Xaa
FT /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
FT N,N,N-trimethyl Lys"
FT Misc-difference 11 /note= "optionally hydroxy-Pro"
FT
XX
XX WO200044769-A1.
PN
XX
XX 03-AUG-2000.
PD
XX
XX 28-JAN-2000; 2000OWO-US01978.
PF
XX
XX 29-JAN-1999; 99US-0118381.
PR
XX 28-DEC-1999; 99US-0173343.
XX
XX (UTAH ) UNIV UTAH RES FOUND.
PA

```


XX McIntosh JM, Olivera BM, Cruz LJ;
XX WPI: 2000-476222/41.
DR
XX
PT Purified ap-conotoxin derived from cone snail venom for use as an
analgesic -
PS Claim 2: Page 19; 29pp; English.
XX
CC The present sequence represents an ap-conotoxin peptide, designated
Mar2. Conotoxins are naturally available in minute amounts in the
venom of cone snails. The peptides have analgesic activity. The
peptides are used to treat or prevent pain.
XX
SQ Sequence 12 AA;
Query Match 74.1%; Score 60; DB 21; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GVCCGKILCHXC 13
Db 1 GVCCGAXLCHPC 12
RESULT 9
AAB08015
ID AAB08015 standard; peptide; 12 AA.
XX
AC AAB08015;
XX
DT 14-NOV-2000 (first entry)
XX
DE Generic formula for conotoxin peptide U036.
XX
KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; U036.
XX
OS Conus marmoreus.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /label= Tyr, Xaa
FT /note= "Xaa is mono-halo-Tyr, di-halo-Tyr,
FT O-sulfo-Tyr, O-phospho Tyr or nitro-Tyr"
FT Misc-difference 7 /label= Lys, Xaa
FT /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
FT N,N,N-trimethyl Lys"
FT Misc-difference 11 /note= "optionally hydroxy-Pro"
FT
FT
FN WO200044769-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US01978.
XX
PR 29-JAN-1999; 99US-0118381.
PR 28-DEC-1999; 99US-0173343.
XX
XX (UTAH) UNIV UTAH RES FOUND.
XX
PI McIntosh JM, Olivera BM, Cruz LJ;
XX
DR WPI: 2000-476222/41.
XX
XX Purified ap-conotoxin derived from cone snail venom for use as an
analgesic -
PS Claim 2: Page 19; 29pp; English.
XX
CC The present sequence represents an ap-conotoxin peptide, designated

CC U036. Conotoxins are naturally available in minute amounts in the
venom of cone snails. The peptides have analgesic activity. The
peptides are used to treat or prevent pain.
XX
SQ Sequence 12 AA;
Query Match 74.1%; Score 60; DB 21; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GVCCGKILCHXC 13
Db 1 GVCCGAXLCHPC 12
RESULT 10
AAB08013
ID AAB08013 standard; peptide; 13 AA.
XX
AC AAB08013;
XX
DT 14-NOV-2000 (first entry)
XX
DE Generic formula for conotoxin peptide Mar1.
XX
KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Mar1.
XX
OS Conus marmoreus.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 7 /label= Tyr, Xaa
FT /note= "Xaa is mono-halo-Tyr, di-halo-Tyr,
FT O-sulfo-Tyr, O-phospho Tyr or nitro-Tyr"
FT Misc-difference 8 /label= Lys, Xaa
FT /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
FT N,N,N-trimethyl Lys"
FT Misc-difference 12 /note= "optionally hydroxy-Pro"
FT
FT
FN WO200044769-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US01978.
XX
PR 29-JAN-1999; 99US-0118381.
PR 28-DEC-1999; 99US-0173343.
XX
XX (UTAH) UNIV UTAH RES FOUND.
XX
PI McIntosh JM, Olivera BM, Cruz LJ;
XX
DR WPI: 2000-476222/41.
XX
XX Purified ap-conotoxin derived from cone snail venom for use as an
analgesic -
PS Claim 2: Page 19; 29pp; English.
XX
XX The present sequence represents an ap-conotoxin peptide, designated
Mar1. Conotoxins are naturally available in minute amounts in the
venom of cone snails. The peptides have analgesic activity. The
peptides are used to treat or prevent pain.
XX
SQ Sequence 13 AA;
Query Match 74.1%; Score 60; DB 21; Length 13;
Best Local Similarity 75.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GVCCGKILCHXC 13

Db 2 GVCGGXKLCHPC 13

RESULT 11

AA07664 standard; protein; 174 AA.

AA07664;

13-FEB-1991 (first entry)

Cysteine-added variant of Granulocyte-Colony Stimulating Factor.

cysteine-added variant; cancer chemotherapy;

granulocyte-colony stimulating factor (G-CSF).

Homo sapiens.

MO9012874-A.

01-NOV-1990.

19-APR-1990; 90WO-US02144.

21-APR-1989; 89US-0341990.

(GENE-) GENETICS INST INC.

Shaw G, Veldman G, Woeters J;

WPI: 1990-348477/46.

Cysteine-added variants of interleukin-3, erythropoietin etc.

contain at least 1 cysteine residue attached to polyalkylene

glycol gp. for improved pharmacokinetic properties

Disclosure; Page 15; 46pp; English.

Ala at position 37 of mature, human G-CSF is replaced by Cys. The

Cys residue is attached to a polyalkylene glycol moiety

which enhances the pharmacokinetic properties of the protein. In

addition, Cys at position 17 is replaced by Ala to prevent possible

improper disulphide bridge formation. The Thr and Pro residues at

positions 1 and 2 of native G-CSF (= positions 2 and 3 of this

sequence) may also be deleted to permit more consistent removal

of the N-terminal methionine.

See also AA07663, AA07665 and AA08358-9.

Sequence 174 AA;

Query Match 61.7%; Score 50; DB 11; Length 174;

Best Local Similarity 77.8%; Pred. No. 20;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

3 VCGGKYLCH 11

35 LCCTYKLC 43

RESULT 12

ABG15487 standard; Protein; 369 AA.

ABG15487;

18-FEB-2002 (first entry)

Novel human diagnostic protein #15478.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

Food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

MO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dremanac RT, Liu C, Tang YT;

WPI: 2001-639362/73.

N-PSDB; AAS79674.

New isolated polynucleotide and encoded polypeptides, useful in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess

biodiversity

Claim 20: SEQ ID NO 45846; 103pp; English.

The invention relates to isolated polynucleotide (I) and

polypeptide (II) sequences. (I) is useful as hybridisation probes,

polymerase chain reaction (PCR) primers, oligomers, and for chromosome

and gene mapping, and in recombinant production of (II). The

polynucleotides are also used in diagnostics as expressed sequence tags

for identifying expressed genes. (I) is useful in gene therapy techniques

to restore normal activity of (II) or to treat disease states involving

(II). (II) is useful for generating antibodies against it, detecting or

quantitating a polypeptide in tissue, as molecular weight markers and as

a food supplement. (II) and its binding partners are useful in medical

imaging of sites expressing (II). (I) and (II) are useful for treating

disorders involving aberrant protein expression or biological activity.

The polypeptide and polynucleotide sequences have applications in

diagnostics, forensics, gene mapping, identification of mutations

and to produce other types of data and products dependent on DNA and

amino acid sequences. ABG00010-ABG30377 represent novel human

diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed

specification, but was obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 369 AA;

Query Match 58.0%; Score 47; DB 22; Length 369;

Best Local Similarity 46.2%; Pred. No. 92;

Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

1 VCGGKYLCHXC 13

56 VTTCGSHFRCPC 68

RESULT 13

ABB66559 standard; Protein; 55 AA.

ABB66559;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 26469.

Drosophila; developmental biology; cell signalling; insecticide;

pharmaceutical.

Drosophila melanogaster.

PM WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL10662.
XX
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
PS Disclosure; SEQ ID NO 26469; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 55 AA:

Query Match 56.8%; Score 46; DB 22; Length 55;
Best Local Similarity 58.3%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYCCGYKLCCHXC 13
Db 7 GSCCGYCCGCP 18

RESULT 14
AA83941
ID AA83941 standard; Protein; 69 AA.
XX
AC AA83941;
XX
XX 28-JUL-2000 (first entry)
DT
DE Benzene ring hydroxylase ferredoxin subunit.
XX
XX Oxygenase; hydroxyl; benzene ring hydroxylase; hydroxylation;
KM ferredoxin-NAD(P)H oxidoreductase; enzyme; phenanthrene; anthracene.
XX
OS Nocardioides sp. strain KP7.
XX
PN JP2000069979-A.
XX
PD 07-MAR-2000.
XX
PF 28-AUG-1998; 98JP-0259413.
XX
PR 28-AUG-1998; 98JP-0259413.
XX
PA (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX
XX WPI; 2000-353293/31.
DR N-PSDB; AAA10035.
XX

PT A gene group encoding an enzyme, introduces hydroxyl group to the
PT benzene ring of a polycyclic aromatic compound -
XX
XX
PS Claim 3; Page 10-11; 16pp; Japanese.
XX
CC The invention relates to the isolation and cloning of gene encoding
CC subunits of an oxygenase which introduces a hydroxyl group to a benzene
CC ring (a benzene ring hydroxylase). The complex comprises alpha and beta
CC subunits of the oxygenase (AA83939 and AA83940), a ferredoxin subunit
CC (AA83941) and a ferredoxin-NAD(P)H oxidoreductase subunit (AA83942).
CC This sequence corresponds to the ferredoxin subunit of the complex. The
CC complex prepared from the encoding genes is useful as an enzyme reagent
CC for hydroxylating phenanthrene and anthracene.
XX
SQ Sequence 69 AA:

Query Match 56.8%; Score 46; DB 21; Length 69;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 CCGYKLC 10
Db 10 CCGYKLC 16

RESULT 15
AA06111
ID AA06111 standard; Protein; 538 AA.
XX
XX AA06111;
AC
XX
XX 24-OCT-2001 (first entry)
DT
XX
DE Novel human polypeptide PROTX.
XX
XX Human; PROTX; PROTX1; AJUBA-like; central nervous system disorder; CNS;
KM cancer; neuromuscular disorder; cardiac disorder; clone AL132780A.
XX
OS Homo sapiens.
XX
PM WO200158946-A2.
XX
PD 16-AUG-2001.
XX
PF 08-FEB-2001; 2001WO-US04402.
XX
PR 08-FEB-2000; 2000US-0180880.
PR 08-FEB-2000; 2000US-0181044.
PR 10-FEB-2000; 2000US-0181656.
PR 15-FEB-2000; 2000US-0182795.
PR 07-FEB-2001; 2001US-0182795.
XX
PA (CURA-) CURAGEN CORP.
XX
XX
PI Taupier RJ, Majumder K, Vernet CM, Prayaga SK;
XX
XX WPI; 2001-488970/53.
DR N-PSDB; AAS09145.
XX
XX
PT Isolated AJUBA-like, keratin-like and endothelin polypeptides for the
PT treatment and diagnosis of cancers and neurological disorders -
XX
XX
PS Claim 1; Page 9-10; 14pp; English.
XX
CC The present invention relates to the isolation of 4 novel human
CC polypeptides, termed PROTX polypeptides (AA06111-AA06114). The
CC PROTX polypeptides are AJUBA-like, keratin-like or endothelin
CC polypeptides. Polynucleotides encoding PROTX, PROTX polypeptides and
CC antibodies that bind PROTX polypeptides are useful for treating or
CC preventing a pathology associated with altered levels of PROTX especially
CC in a human subject. PROTX polypeptides, PROTX polynucleotides and PROTX
CC antibodies are useful for identifying an agent that binds to PROTX, for
CC identifying potential therapeutic agents for use in a treatment of a

CC pathology related to aberrant expression or physiological interaction
 CC with PROTX, for screening for a modulator of activity of PROTX and
 CC for determining the presence of or a predisposition to a disease
 CC associated with altered levels of PROTX. PROTX, an AUBA-like protein,
 CC is useful for treating central nervous system disorders, cancer,
 CC Parkinson's disease, Alzheimer's disease, neuromuscular and cardiac
 CC disorders. PROTX, a keratin, is useful for treating inflammatory and
 CC neoplastic pancreatic disorders. Protx, an endothelin, is useful for
 CC treating hypertension, cardiovascular and neurological disorders. The
 CC present sequence represents human polypeptide PROTX. The DNA sequence
 CC encoding PROTX is isolated from the genomic clone AL132780_A.

XX Sequence 538 AA:

Query Match 56.8%; Score 46; DB 22; Length 538;

Best Local Similarity 50.0%; Pred. No. 1.7e+02;

Matches 8; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 2 GVCC---GYKLCHXC 13

DB 508 GCCCFPLDGDHLLCHGC 523

Search completed: December 2, 2002, 10:07:07
 Job time : 31.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:06:06 ; Search time 12 seconds
(without alignments)
31.875 Million cell updates/sec

Title: US-09-787-986A-2

Perfect score: 81

Sequence: 1 VGVCCGYKLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/CTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfill1es1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	55.6	485	2	US-08-724-394A-8
2	43	53.1	574	1	US-08-049-473-2
3	43	53.1	574	1	US-08-312-648-2
4	43	53.1	574	5	PCR-US94-04190-2
5	42	51.9	177	2	US-08-431-459A-30
6	42	51.9	177	3	US-09-122-443-8
7	42	51.9	646	4	US-09-625-188-10
8	41	50.6	171	3	US-08-833-167-44
9	41	50.6	171	3	US-08-833-167-45
10	41	50.6	171	3	US-08-833-167-47
11	41	50.6	171	3	US-08-833-167-57
12	41	50.6	171	3	US-08-833-167-59
13	41	50.6	171	3	US-08-833-167-60
14	41	50.6	171	4	US-09-344-837A-44
15	41	50.6	171	4	US-09-344-837A-45
16	41	50.6	171	4	US-09-344-837A-47
17	41	50.6	171	4	US-09-344-837A-57
18	41	50.6	171	4	US-09-344-837A-59
19	41	50.6	171	4	US-09-344-837A-60
20	41	50.6	173	1	US-08-354-456A-7
21	41	50.6	174	1	US-08-010-099-82
22	41	50.6	174	1	US-08-225-224-5
23	41	50.6	174	1	US-08-434-411-2
24	41	50.6	174	1	US-08-434-407-2
25	41	50.6	174	1	US-08-448-716-82
26	41	50.6	174	1	US-08-783-288-2
27	41	50.6	174	2	US-08-431-459A-31

28	41	50.6	174	2	US-08-890-640-2	Sequence 2, Appli
29	41	50.6	174	3	US-08-722-258-5	Sequence 5, Appli
30	41	50.6	174	3	US-08-833-167-49	Sequence 49, Appli
31	41	50.6	174	3	US-08-833-167-50	Sequence 50, Appli
32	41	50.6	174	3	US-08-833-167-51	Sequence 51, Appli
33	41	50.6	174	3	US-08-833-167-52	Sequence 52, Appli
34	41	50.6	174	3	US-08-833-167-53	Sequence 53, Appli
35	41	50.6	174	3	US-08-833-167-54	Sequence 54, Appli
36	41	50.6	174	3	US-08-833-167-55	Sequence 55, Appli
37	41	50.6	174	3	US-08-833-167-56	Sequence 56, Appli
38	41	50.6	174	3	US-08-833-167-95	Sequence 95, Appli
39	41	50.6	174	3	US-08-833-167-96	Sequence 96, Appli
40	41	50.6	174	3	US-08-833-167-97	Sequence 97, Appli
41	41	50.6	174	3	US-08-833-167-98	Sequence 98, Appli
42	41	50.6	174	3	US-08-833-167-99	Sequence 99, Appli
43	41	50.6	174	3	US-08-833-167-100	Sequence 100, App
44	41	50.6	174	3	US-08-833-167-101	Sequence 101, App
45	41	50.6	174	3	US-08-833-167-102	Sequence 102, App

ALIGNMENTS

RESULT 1
US-08-724-394A-8
; Sequence 8, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..485
; OTHER INFORMATION: /note="Roret"
; US-08-724-394A-8

Query Match 55.6%; Score 45; DB 2; Length 485;
Best Local Similarity 46.2%; Pred. No. 55;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 VGVCGYKLCXHC 13
I : : : : :
DB 27 VSLNGHSHYCHLC 39

RESULT 2
US-08-049-473-2
Sequence 2, Application US/08049473
Patent No. 5386021

GENERAL INFORMATION:

APPLICANT: Moss, Joel
APPLICANT: Mishima, Koichi
APPLICANT: Nightingale, Maria
APPLICANT: Tsuchiya, Mikako
TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
TITLE OF INVENTION: PROTEIN WITH AN ADP-RIBOSYLATION FACTOR DOMAIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/049,473
FILING DATE: 19930419
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael L.
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH050.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 574 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-049-473-2

Query Match 53.1%; Score 43; DB 1; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 16; Gaps 1;

OY 2 GVC-----CGYKLCXHC 13
I : : : : :
DB 32 GVCEDVFSLQGDKVPRLLCGHTVCHDC 59

RESULT 3
US-08-312-648-2
Sequence 2, Application US/08312648
Patent No. 5514600

GENERAL INFORMATION:

APPLICANT: Moss, Joel
APPLICANT: Mishima, Koichi
APPLICANT: Nightingale, Maria
APPLICANT: Tsuchiya, Mikako
TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
TITLE OF INVENTION: PROTEIN WITH AN ADP-RIBOSYLATION FACTOR DOMAIN

NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,648
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,473
FILING DATE: 19-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael L.
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH050.001DV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 574 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-312-648-2

Query Match 53.1%; Score 43; DB 1; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 16; Gaps 1;

OY 2 GVC-----CGYKLCXHC 13
I : : : : :
DB 32 GVCEDVFSLQGDKVPRLLCGHTVCHDC 59

RESULT 4
PCT-US94-04190-2
Sequence 2, Application PC/TUS9404190

GENERAL INFORMATION:

APPLICANT: The Government of the United States of America
APPLICANT: as represented by the Secretary, Department
APPLICANT: of Health and Human Services
TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
TITLE OF INVENTION: PROTEIN WITH AN ADP-RIBOSYLATION FACTOR DOMAIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04190
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael L.
REGISTRATION NUMBER: 36,516

REFERENCE/DOCKET NUMBER: NIH050.0010PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 574 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-04190-2

Query Match 53.1%; Score 43; DB 5; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 16; Gaps 1;

OY 2 GVC-----CGKLCXHC 13
||| 11:111
Db 32 GVCEDVFSLQGDVPRLLCGHTVCHDC 59

RESULT 5
US-08-431-459A-30
; Sequence 30, Application US/08431459A
; Patent No. 5840543
; GENERAL INFORMATION:
; APPLICANT: Hockney, Robert C.
; APPLICANT: Kara, Bhupendra V.
; TITLE OF INVENTION: FERMENTATION PROCESS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44M diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431.459A
; FILING DATE: 01-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,676
; FILING DATE: 22-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/661,306
; FILING DATE: 27-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9004390.2
; FILING DATE: 27-FEB-1990
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-431-459A-30

Query Match 51.9%; Score 42; DB 2; Length 177;
Best Local Similarity 63.6%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 VGVCCGYKLCH 11
| | | | |
Db 36 VSECATYKLCH 46

RESULT 6
US-09-122-443-8

; Sequence 8, Application US/09122443
; Patent No. 6060284
; GENERAL INFORMATION:
; APPLICANT: Bazar, J. Fernando
; TITLE OF INVENTION: MAMMALIAN CYTOKINE; RELATED REAGENTS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/122.443
; FILING DATE: 24-JUL-1998
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,765
; FILING DATE: 25-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0758K1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-122-443-8

Query Match 51.9%; Score 42; DB 3; Length 177;
Best Local Similarity 63.6%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 VGVCCGYKLCH 11
| | | | |
Db 36 VSECATYKLCH 46

RESULT 7
US-09-625-188-10
; Sequence 10, Application US/09625188
; Patent No. 6307037
; GENERAL INFORMATION:
; APPLICANT: No. 6307037artlis AG
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-31285PI
; CURRENT APPLICATION NUMBER: US/09/625.188
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Ashbya gossypii
US-09-625-188-10

Query Match 51.9%; Score 42; DB 4; Length 646;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13

Db 52 CGYICQFC 60

RESULT 8

US-08-833-167-44
Sequence 44, Application US/08833167
Patent No. 6100070

GENERAL INFORMATION:

APPLICANT: ZURELUH, LINDA L.
APPLICANT: MCWHERTER, CHARLES A
APPLICANT: MCKEARN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIJOING
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,167
FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:

NAME: BENNETT, DENNIS A
REFERENCE/DOCKET NUMBER: 2907/1
TELECOMMUNICATION INFORMATION:

TELEPHONE: 314-694-5402
TELEFAX: 314-694-9095

INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-833-167-44

Query Match 50.6%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 78;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCH 11

Db 110 LCATYKLCH 118

RESULT 9

US-08-833-167-45
Sequence 45, Application US/08833167
Patent No. 6100070

GENERAL INFORMATION:

APPLICANT: ZURELUH, LINDA L.
APPLICANT: MCWHERTER, CHARLES A
APPLICANT: MCKEARN, JOHN P

APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIJOING
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,167
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENNETT, DENNIS A
REFERENCE/DOCKET NUMBER: 2907/1
TELECOMMUNICATION INFORMATION:

TELEPHONE: 314-694-5402
TELEFAX: 314-694-9095

INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-833-167-45

Query Match 50.6%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 78;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCH 11

Db 74 LCATYKLCH 82

RESULT 10

US-08-833-167-47
Sequence 47, Application US/08833167
Patent No. 6100070

GENERAL INFORMATION:

APPLICANT: ZURELUH, LINDA L.
APPLICANT: MCWHERTER, CHARLES A
APPLICANT: MCKEARN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIJOING
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS

COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,167
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENNETT, DENNIS A
REFERENCE/DOCKET NUMBER: 2907/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-5402
TELEFAX: 314-694-9095
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-167-47

Query Match 50.6%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VCCGYKLCH 11
: | |||||
Db 65 LCATYKLC 73

RESULT 11
US-08-833-167-57
Sequence 57, Application US/08833167
Patent No. 6100070
GENERAL INFORMATION:
APPLICANT: ZURELUH, LINDA L
APPLICANT: MCWHERTER, CHARLES A
APPLICANT: MCKEARN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIONG
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
ADDRESS: CORPORATE PATENT DEPT.,
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,167
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENNETT, DENNIS A
REFERENCE/DOCKET NUMBER: 2907/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-5402
TELEFAX: 314-694-9095
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-167-57

Query Match 50.6%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VCCGYKLCH 11
: | |||||
Db 110 LCATYKLC 118

RESULT 12
US-08-833-167-59
Sequence 59, Application US/08833167
Patent No. 6100070
GENERAL INFORMATION:
APPLICANT: ZURELUH, LINDA L
APPLICANT: MCWHERTER, CHARLES A
APPLICANT: MCKEARN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIONG
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
ADDRESS: CORPORATE PATENT DEPT.,
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,167
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENNETT, DENNIS A
REFERENCE/DOCKET NUMBER: 2907/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-5402
TELEFAX: 314-694-9095
INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-167-59

Query Match 50.6%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGKYLCH 11
: | |||||
Db 74 LCATYKILCH 82

RESULT 13
US-08-833-167-60
; Sequence 60, Application US/08833167
; Patent No. 6100070
; GENERAL INFORMATION:
; APPLICANT: ZURELUH, LINDA L
; APPLICANT: MCWHERTER, CHARLES A
; APPLICANT: MCKEARN, JOHN P
; APPLICANT: KLEIN, BARBARA K
; APPLICANT: FENG, YI QING
; APPLICANT: BRAFORD-GOLDBERG, SARAH R
; APPLICANT: LEE, STEPHEN C
; TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
; NUMBER OF SEQUENCES: 129
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DENNIS A BENNETT, G.D. SEARLE & CO.,
; STREET: P.O. BOX 5110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,167
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US 96/15935
; FILING DATE: 04-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,382
; FILING DATE: 05-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BENNETT, DENNIS A
; REFERENCE/DOCKET NUMBER: 2907/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-694-5402
; TELEFAX: 314-694-9095
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-833-167-60

Query Match 50.6%; Score 41; DB 3; Length 171;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGKYLCH 11
: | |||||
Db 65 LCATYKILCH 73

RESULT 14
US-09-344-837A-44
; Sequence 44, Application US/09344837A
; Patent No. 6358505
; GENERAL INFORMATION:
; APPLICANT: ZURELUH, LINDA L
; APPLICANT: MCWHERTER, CHARLES A
; APPLICANT: MCKEARN, JOHN P
; APPLICANT: KLEIN, BARBARA K
; APPLICANT: FENG, YI QING
; APPLICANT: BRAFORD-GOLDBERG, SARAH R
; APPLICANT: LEE, STEPHEN C
; TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
; NUMBER OF SEQUENCES: 129
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: S. CHRISTOPHER BAUER
; ADDRESSEE: MONSANTO/G. D. SEARLE & CO.,
; ADDRESSEE: PATENT DEPARTMENT CENTRAL
; STREET: P.O. BOX 5110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,837A
; FILING DATE: 25-JUN-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US 96/15935
; FILING DATE: 04-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,382
; FILING DATE: 05-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: S. CHRISTOPHER BAUER
; REFERENCE/DOCKET NUMBER: 2907/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 636-737-6257
; TELEFAX: 636-737-5452
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-344-837A-44

Query Match 50.6%; Score 41; DB 4; Length 171;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGKYLCH 11
: | |||||
Db 110 LCATYKILCH 118

RESULT 15
US-09-344-837A-45
; Sequence 45, Application US/09344837A
; Patent No. 6358505
; GENERAL INFORMATION:
; APPLICANT: ZURELUH, LINDA L

APPLICANT: MCWHERTER, CHARLES A
APPLICANT: MCKEARN, JOHN P
APPLICANT: KLEIN, BARBARA K
APPLICANT: FENG, YIOLING
APPLICANT: BRAFORD-GOLDBERG, SARAH R
APPLICANT: LEE, STEPHEN C
TITLE OF INVENTION: G-CSF RECEPTOR AGONISTS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: S. CHRISTOPHER BAUER
ADDRESSEE: MONSANTO/G. D. SEARLE & CO.,
ADDRESSEE: PATENT DEPARTMENT CENTRAL
STREET: P.O. BOX 5110
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,837A
FILING DATE: 25-JUN-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 96/15935
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,382
FILING DATE: 05-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: S. CHRISTOPHER BAUER
REFERENCE/DOCKET NUMBER: 2907/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 636-737-6257
TELEFAX: 636-737-5452
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-344-837A-45

Query Match 50.6%; Score 41; DB 4; Length 171;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCGGYKLCH 11
DB 74 LCATYKRLCH 82

Search completed: December 2, 2002, 10:09:27
Job time : 12 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:07:11 : Search time 7.5 Seconds
(without alignments)
27.602 Million cell updates/sec

Title: US-09-787-986a-2
Perfect score: 81
Sequence: 1 VGVCCGYKLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications - A1:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	56.8	538	10	US-09-779-307-2
2	46	56.8	538	10	US-09-779-307-11
3	46	56.8	547	10	US-09-779-307-12
4	46	56.8	547	10	US-09-779-307-13
5	43	53.1	26	8	US-08-775-765-9
6	42	51.9	24	8	US-08-775-765-8
7	41	50.6	174	10	US-09-754-532-82
8	41	50.6	174	10	US-09-760-008A-1
9	41	50.6	174	10	US-09-950-473-2
10	41	50.6	174	12	US-10-016-403-1
11	41	50.6	174	12	US-10-016-403-2
12	41	50.6	174	12	US-10-016-403-3
13	41	50.6	174	12	US-10-016-403-4
14	41	50.6	174	12	US-10-003-496-1
15	41	50.6	175	10	US-09-754-532-67
16	41	50.6	175	10	US-09-754-532-67
17	41	50.6	175	10	US-09-754-532-68
18	41	50.6	175	10	US-09-754-532-69
19	41	50.6	175	10	US-09-754-532-71

20	41	50.6	175	10	US-09-754-532-76	Sequence 76, Appl
21	41	50.6	175	10	US-09-754-532-78	Sequence 78, Appl
22	41	50.6	175	10	US-09-754-532-79	Sequence 79, Appl
23	41	50.6	175	10	US-09-754-532-80	Sequence 80, Appl
24	41	50.6	175	10	US-09-754-532-81	Sequence 81, Appl
25	41	50.6	175	10	US-09-754-532-85	Sequence 85, Appl
26	41	50.6	175	10	US-09-754-532-86	Sequence 86, Appl
27	41	50.6	175	10	US-09-754-532-87	Sequence 87, Appl
28	41	50.6	175	10	US-09-754-532-88	Sequence 88, Appl
29	41	50.6	175	10	US-09-754-532-89	Sequence 89, Appl
30	41	50.6	175	10	US-09-754-532-90	Sequence 90, Appl
31	41	50.6	175	10	US-09-754-532-91	Sequence 91, Appl
32	41	50.6	175	10	US-09-754-532-92	Sequence 92, Appl
33	41	50.6	175	10	US-09-754-532-93	Sequence 93, Appl
34	41	50.6	175	10	US-09-754-532-94	Sequence 94, Appl
35	41	50.6	175	10	US-09-754-532-95	Sequence 95, Appl
36	41	50.6	175	10	US-09-754-532-96	Sequence 96, Appl
37	41	50.6	175	10	US-09-754-532-97	Sequence 97, Appl
38	41	50.6	175	10	US-09-754-532-98	Sequence 98, Appl
39	41	50.6	175	10	US-09-754-532-99	Sequence 99, Appl
40	41	50.6	175	10	US-09-754-532-100	Sequence 100, Appl
41	41	50.6	175	10	US-09-754-532-101	Sequence 101, Appl
42	41	50.6	175	10	US-09-754-532-102	Sequence 102, Appl
43	41	50.6	175	10	US-09-754-532-104	Sequence 104, Appl
44	41	50.6	175	10	US-09-754-532-105	Sequence 105, Appl
45	41	50.6	175	10	US-09-754-532-106	Sequence 106, Appl

ALIGNMENTS

RESULT 1
US-09-779-307-2
; Sequence 2, Application US/09779307
; Patent No. US20020137675A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-662 US
; CURRENT APPLICATION NUMBER: US/09/779,307
; CURRENT FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-307-2

Query Match 56.8%; Score 46; DB 10; Length 538;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 8; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 2 GVCC---GYKLCHXC 13
| | | | |
Db 508 GCCCFPLDGHILCHXC 523

RESULT 2
US-09-779-307-11
; Sequence 11, Application US/09779307
; Patent No. US20020137675A1

GENERAL INFORMATION:
APPLICANT: Taupier Jr., Raymond
APPLICANT: Majumder, Kunnud
APPLICANT: Vernet, Corine
APPLICANT: Prayaga, Sudhirdas
TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
FILE REFERENCE: 15966-662 US
CURRENT APPLICATION NUMBER: US/09/779,307
CURRENT FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/180,880
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/181,044
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/181,656
PRIOR FILING DATE: 2000-02-10
PRIOR APPLICATION NUMBER: 60/182,795
PRIOR FILING DATE: 2000-02-15
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 538
TYPE: PRT
ORGANISM: Homo sapiens
US-09-779-307-11

Query Match 56.8%; Score 46; DB 10; Length 538;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 8; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

OY 2 GVCC---GYKLCHXC 13
1 11 1: 111 1
Db 508 GCCCFPLDGHILCHGC 523

RESULT 3
US-09-779-307-12
Sequence 12, Application US/09779307
Patent No. US20020137675A1
GENERAL INFORMATION:
APPLICANT: Taupier Jr., Raymond
APPLICANT: Majumder, Kunnud
APPLICANT: Vernet, Corine
APPLICANT: Prayaga, Sudhirdas
TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
FILE REFERENCE: 15966-662 US
CURRENT APPLICATION NUMBER: US/09/779,307
CURRENT FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/180,880
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/181,044
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/181,656
PRIOR FILING DATE: 2000-02-10
PRIOR APPLICATION NUMBER: 60/182,795
PRIOR FILING DATE: 2000-02-15
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 547
TYPE: PRT
ORGANISM: Mus musculus
US-09-779-307-12

Query Match 56.8%; Score 46; DB 10; Length 547;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

OY 2 GVCC---GYKLCHXC 13
1 11 1: 111 1
Db 517 GCCCFPLDGHILCHGC 532

RESULT 4

US-09-779-307-13
Sequence 13, Application US/09779307
Patent No. US20020137675A1
GENERAL INFORMATION:
APPLICANT: Taupier Jr., Raymond
APPLICANT: Majumder, Kunnud
APPLICANT: Vernet, Corine
APPLICANT: Prayaga, Sudhirdas
TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
FILE REFERENCE: 15966-662 US
CURRENT APPLICATION NUMBER: US/09/779,307
CURRENT FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/180,880
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/181,044
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/181,656
PRIOR FILING DATE: 2000-02-10
PRIOR APPLICATION NUMBER: 60/182,795
PRIOR FILING DATE: 2000-02-15
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 547
TYPE: PRT
ORGANISM: Mus musculus
US-09-779-307-13

Query Match 56.8%; Score 46; DB 10; Length 547;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

OY 2 GVCC---GYKLCHXC 13
1 11 1: 111 1
Db 517 GCCCFPLDGHILCHGC 532

RESULT 5
US-08-775-765-9
Sequence 9, Application US/08775765C
Patent No. US20010010821A1
GENERAL INFORMATION:
APPLICANT: Kelleher, Dermot
APPLICANT: Windle, Henry
APPLICANT: Byrne, William
APPLICANT: McManus, Ross
TITLE OF INVENTION: Helicobacter Proteins and Vaccines
FILE REFERENCE: 08/775 765
CURRENT APPLICATION NUMBER: US/08/775,765C
CURRENT FILING DATE: 1996-12-31
EARLIER APPLICATION NUMBER: IE 94 0538
EARLIER FILING DATE: 1994-07-01
EARLIER APPLICATION NUMBER: IE 95 0249
EARLIER FILING DATE: 1995-04-06
EARLIER APPLICATION NUMBER: PCT/IE 95/00036
EARLIER FILING DATE: 1995-07-03
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 26
TYPE: PRT
ORGANISM: Helicobacter pylori
US-08-775-765-9

Query Match 53.1%; Score 43; DB 8; Length 26;
Best Local Similarity 66.7%; Pred. No. 6.8;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 GVCCGYKLC 10
1 1111 1
Db 8 GCCCGYKLC 16

```
RESULT 6
US-08-775-765-8
; Sequence 8, Application US/08775765C
; Patent No. US20010010821A1
; GENERAL INFORMATION:
; APPLICANT: Kelleher, Dermot
; APPLICANT: Windle, Henry
; APPLICANT: Byrne, William
; APPLICANT: McManus, Ross
; TITLE OF INVENTION: Helicobacter Proteins and Vaccines
; FILE REFERENCE: 08/775 765
; CURRENT APPLICATION NUMBER: US/08/775,765C
; CURRENT FILING DATE: 1996-12-31
; EARLIER APPLICATION NUMBER: IE 94 0538
; EARLIER FILING DATE: 1994-07-01
; EARLIER APPLICATION NUMBER: IE 95 0249
; EARLIER FILING DATE: 1995-04-06
; EARLIER APPLICATION NUMBER: PCT/IE 95/00036
; EARLIER FILING DATE: 1995-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-775-765-8

Query Match          51.9%; Score 42; DB 8; Length 24;
Best Local Similarity 50.0%; Pred. No. 8.4;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 GVCCKKLCXHC 13
DB 6 GGCGCTTCCTC 17

RESULT 7
US-09-754-532-82
; Sequence 82, Application US/09754532
; Patent No. US20010016191A1
; GENERAL INFORMATION:
; APPLICANT: Osslund, Timothy D.
; TITLE OF INVENTION: G-CSF ANALOG COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: Amgen Center, 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States of America
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/754,532
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/448,716
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pessin, Karol
; REGISTRATION NUMBER: 34,899
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 805/499-5725
; TELEFAX: 805/499-8011
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-754-532-82

Query Match          50.6%; Score 41; DB 10; Length 174;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCX 11
DB 36 LCATYKLCX 44

RESULT 8
US-09-760-008A-1
; Sequence 1, Application US/09760008A
; Patent No. US20020004483A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
; APPLICANT: MIKKELSEN, JAN MOLLER
; TITLE OF INVENTION: G-CSF COMPOUNDS
; FILE REFERENCE: 31-000700US
; CURRENT APPLICATION NUMBER: US/09/760,008A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: DK PA 2000 00341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-008A-1

Query Match          50.6%; Score 41; DB 10; Length 174;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCX 11
DB 35 LCATYKLCX 43

RESULT 9
US-09-950-473-2
; Sequence 2, Application US/09950473
; Patent No. US20020151488A1
; GENERAL INFORMATION:
; APPLICANT: Sarkar, Casim
; APPLICANT: Lauffenburger, Douglas
; TITLE OF INVENTION: G-CSF Analog Compositions and Methods
; FILE REFERENCE: 01017/37732
; CURRENT APPLICATION NUMBER: US/09/950,473
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-473-2
```

Query Match 50.6%; Score 41; DB 10; Length 174;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKLCH 11
: 1 11111
DB 35 LCATYKLCH 43

RESULT 10
US-10-016-403-1
; Sequence 1, Application US/10016403
; Patent No. US20020107505A1
; GENERAL INFORMATION:
; APPLICANT: Holladay, Leslie A.
; TITLE OF INVENTION: MODIFICATION OF POLYPEPTIDE DRUGS TO
; INCREASE ELECTROTRANSPORT FLUX
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Stroud, Stroud, Willink, Thompson & Howard
; STREET: 25 West Main Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2236
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/016,403
; FILING DATE: 10-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/466,610
; FILING DATE: 1995-JUN-06
; ATTORNEY/AGENT INFORMATION:
; NAME: Frenchick, Grady J.
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 8734.28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-257-2281
; TELEFAX: 608-257-7643
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..174
; OTHER INFORMATION: /note="granulocyte-colony
; stimulating factor"
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-016-403-1
Query Match 50.6%; Score 41; DB 12; Length 174;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 VCCGYKLCH 11
: 1 11111
DB 35 LCATYKLCH 43
RESULT 11
US-10-016-403-2
; Sequence 2, Application US/10016403
; Patent No. US20020107505A1
; GENERAL INFORMATION:
; APPLICANT: Holladay, Leslie A.

; TITLE OF INVENTION: MODIFICATION OF POLYPEPTIDE DRUGS TO
; INCREASE ELECTROTRANSPORT FLUX
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Stroud, Stroud, Willink, Thompson & Howard
; STREET: 25 West Main Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2236
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/016,403
; FILING DATE: 10-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/466,610
; FILING DATE: 1995-JUN-06
; ATTORNEY/AGENT INFORMATION:
; NAME: Frenchick, Grady J.
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 8734.28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-257-2281
; TELEFAX: 608-257-7643
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..174
; OTHER INFORMATION: /note="modified granulocyte-colony
; stimulating factor"
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-016-403-2
Query Match 50.6%; Score 41; DB 12; Length 174;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 VCCGYKLCH 11
: 1 11111
DB 35 LCATYKLCH 43
RESULT 12
US-10-016-403-3
; Sequence 3, Application US/10016403
; Patent No. US20020107505A1
; GENERAL INFORMATION:
; APPLICANT: Holladay, Leslie A.
; TITLE OF INVENTION: MODIFICATION OF POLYPEPTIDE DRUGS TO
; INCREASE ELECTROTRANSPORT FLUX
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Stroud, Stroud, Willink, Thompson & Howard
; STREET: 25 West Main Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2236
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/016,403
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/466,610
FILING DATE: 1995-JUN-06
ATTORNEY/AGENT INFORMATION:
NAME: Frenchick, Grady J.
REGISTRATION NUMBER: 29,018
REFERENCE/DOCKET NUMBER: 8734.28
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-257-2281
TELEFAX: 608-257-7643
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..174
OTHER INFORMATION: /note= "modified granulocyte-colony
stimulating factor."
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-016-403-3
Query Match 50.6%; Score 41; DB 12; Length 174;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 VCCGYKLCH 11
: | | | | |
DB 35 LCATYKLCH 43
RESULT 13
US-10-016-403-4
Sequence 4, Application US/10016403
Patent No. US20020107505A1
GENERAL INFORMATION:
APPLICANT: Holladay, Leslie A.
TITLE OF INVENTION: MODIFICATION OF POLYPEPTIDE DRUGS TO
INCREASE ELECTROTRANSPORT FLUX
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stroud, Stroud, Willink, Thompson & Howard
STREET: 25 West Main Street
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2236
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/016,403
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/466,610
FILING DATE: 1995-JUN-06
ATTORNEY/AGENT INFORMATION:
NAME: Frenchick, Grady J.
REGISTRATION NUMBER: 29,018
REFERENCE/DOCKET NUMBER: 8734.28
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-257-2281
TELEFAX: 608-257-7643
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids

TYPE: amino acid
TOPOLOGY: linear
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..174
OTHER INFORMATION: /note= "granulocyte-colony
stimulating factor."
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-016-403-4
Query Match 50.6%; Score 41; DB 12; Length 174;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 VCCGYKLCH 11
: | | | | |
DB 35 LCATYKLCH 43
RESULT 14
US-10-003-496-1
Sequence 1, Application US/10003496
Patent No. US20020142964A1
GENERAL INFORMATION:
APPLICANT: Maxygen Aps
APPLICANT: Maxygen Holdings Ltd.
TITLE OF INVENTION: Single-Chain Polypeptides
FILE REFERENCE: 0218us210
CURRENT APPLICATION NUMBER: US/10/003,496
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/245,727
PRIOR FILING DATE: 2000-11-02
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 174
TYPE: PRT
ORGANISM: Homo sapiens
US-10-003-496-1
Query Match 50.6%; Score 41; DB 12; Length 174;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 VCCGYKLCH 11
: | | | | |
DB 35 LCATYKLCH 43
RESULT 15
US-09-754-532-2
Sequence 2, Application US/09754532
Patent No. US20010016191A1
GENERAL INFORMATION:
APPLICANT: Osslund, Timothy D.
TITLE OF INVENTION: G-CSF ANALOG COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: Amgen Center, 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: United States of America
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/754,532
FILING DATE:
CLASSIFICATION:

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/448,716
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pessin, Karol
 ; REGISTRATION NUMBER: 34,899
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 805/499-5725
 ; TELEFAX: 805/499-8011
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 175 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-754-532-2

Query Match 50.6%; Score 41; DB 10; Length 175;
 Best Local Similarity 66.7%; Pred. NO. 53;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VCCGYKILCH 11
 : | | | | |
 DB 36 LCATYKILCH 44

Search completed: December 2, 2002, 10:09:48
 Job time : 7.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:05:26 ; Search time 13.5 seconds
(without alignments)
92.574 Million cell updates/sec

Title: US-09-787-986a-2
Perfect score: 81
Sequence: 1 VCGCGKRLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	57	70.4	258	2 T30368	probable immediate
2	48	59.3	1322	2 T15689	hypothetical prote
3	47	58.0	243	2 T10407	immediate early pr
4	47	58.0	322	2 T08592	TGACG-motif-bindin
5	47	58.0	326	2 T08591	TGACG-motif bindin
6	47	58.0	373	2 T47488	hypothetical prote
7	46	56.8	55	2 S25774	testis-specific pr
8	45	56.8	79	2 P95210	conserved domain p
9	45	55.6	126	2 T42321	hypothetical prote
10	44	54.3	244	2 T26913	hypothetical prote
11	44	54.3	244	2 T26912	hypothetical prote
12	44	54.3	332	2 T24312	hypothetical prote
13	44	54.3	514	2 T10559	hypothetical prote
14	43	53.1	261	2 G72867	hypothetical prote
15	43	53.1	261	2 T41878	IE-0 Orf141 - Bomb
16	43	53.1	574	2 T27752	hypothetical prote
17	43	53.1	574	2 A46054	GTP-binding protei
18	42	51.9	78	2 T50943	probable ferredoxi
19	42	51.9	171	2 G90687	phosphatidylglycer
20	42	51.9	171	2 C85538	phosphatidylglycer
21	42	51.9	172	2 B64771	phosphatidylglycer
22	42	51.9	207	2 A24573	granulocyte colony
23	42	51.9	211	2 T20590	granulocyte colony
24	42	51.9	222	2 T47487	hypothetical prote
25	42	51.9	350	2 G84647	hypothetical prote
26	42	51.9	373	2 B84647	hypothetical prote
27	42	51.9	400	1 ZBBE14	hypothetical prote
28	42	51.9	489	2 S62474	44.1K zinc-binding
29	42	51.9	587	2 A56015	probable transcrip
					finger protein sig

30	41	50.6	98	2 JC5147	tachyctin precurs
31	41	50.6	104	2 F83370	hydrogen cyanide s
32	41	50.6	171	2 AD0554	phosphatidylglycer
33	41	50.6	204	1 F0HUGL	granulocyte colony
34	41	50.6	208	2 A26496	granulocyte colony
35	41	50.6	214	2 JC5043	granulocyte colony
36	41	50.6	295	2 T20629	hypothetical prote
37	41	50.6	348	2 T47494	hypothetical prote
38	41	50.6	639	2 T33166	hypothetical prote
39	41	50.6	884	2 T18649	hypothetical prote
40	41	50.6	1235	2 Q0BEW4	hypothetical protei
41	40.5	50.0	1181	2 D86157	DNA-binding protei
42	40	49.4	109	2 A72615	hypothetical prote
43	40	49.4	179	2 A83068	hypothetical prote
44	40	49.4	525	2 T10574	hypothetical prote
45	40	49.4	553	2 B88072	protein ZK1240.2 [

ALIGNMENTS

```
RESULT 1
T30368
Probable immediate-early transactivator 0 - Lymantria dispar nuclear polyhedrosis vir
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T30368
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavecek, J.M.; R
Virology 253, 17-34, 1999
A>Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantri
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30368
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-258 <K0Z>
A:Cross-references: EMBL:AF081810; PIDN:AACT0206.1
C:Keywords: immediate-early protein

Query Match          70.4%; Score 57; DB 2; Length 258;
Best Local Similarity 63.6%; Pred. No. 0.59;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 VCGGKRLCHXC 13
DB 208 VCGGRVNCAC 218

RESULT 2
T15689
Hypothetical protein C28G1.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
R:Favell, T.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans.csmid C28G1.
A:Reference number: Z18389
A:Accession: T15689
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1322 <FAV>
A:Cross-references: EMBL:U41026; NID:g1086701; PID:g1086702; PIDN:AAA82350.1; CESP:CE
C:Genetics:
A:Gene: CESP:C28G1.3
A:Introns: 25/1; 131/3; 150/1; 166/3; 180/3; 204/2; 235/3; 344/3; 385/2; 436/3; 483/2
A:Superfamily: RING finger homology
F:810-862/Domain: RING finger homology <RRN>

Query Match          59.3%; Score 48; DB 2; Length 1322;
Best Local Similarity 63.6%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 VCGGKRLCHXC 13
```

Db 1071 VTGHALCHXC 1081

RESULT 3

T10407 Immediate early protein 0 - Oryzias pseudotsugata nuclear polyhedrosis virus
C:Species: Oryzias pseudotsugata nuclear polyhedrosis virus, OPMNPV
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C:Accession: T10407
R:Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohmann, G.F.
Virology 229, 381-399, 1997
A:Title: The sequence of the Oryzias pseudotsugata multinnucleocapsid nuclear polyhedrosis
A:Reference number: 217011; MUID:97271300; PMID:9126251
A:Accession: T10407
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-245 <AHR>
A:Cross-references: EMBL:U75930; NID:g2934903; PID:g1911384

Query Match 58.0%; Score 47; DB 2; Length 245;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 CGYKLCXHC 13
||| | | | |
Db 213 CCQYSLCYAC 222

RESULT 4

T08592 TGACG-motif-binding protein STP2 - soybean
C:Species: Glycine max (soybean)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08592
R:Cheong, Y.H.; Yoo, C.M.; Park, J.M.; Ryu, G.R.; Geokjian, V.H.; Nagao, R.T.; Key, J.L.
submitted to the EMBL Data Library, September 1995
A:Description: STP1 is a novel TGACG-binding factor with a zinc-finger motif and a bzip
A:Reference number: 216445
A:Accession: T08592
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-322 <CHE>
A:Cross-references: EMBL:L28004; NID:g986966; PID:g2934885
A:Experimental source: strain Williams; hypocotyl
C:Genetics:
A:Gene: STP2

Query Match 58.0%; Score 47; DB 2; Length 322;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
||| | | | |
Db 61 CGFPLCHSC 69

RESULT 5

T08591 TGACG-motif binding protein STP1 - soybean
C:Species: Glycine max (soybean)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08591
R:Cheong, Y.H.; Yoo, C.M.; Park, J.M.; Ryu, G.R.; Geokjian, V.H.; Nagao, R.T.; Key, J.L.
submitted to the EMBL Data Library, September 1995
A:Description: STP1 is a novel TGACG-binding factor with a zinc-finger motif and a bzip
A:Reference number: 216445
A:Accession: T08591
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-326 <CHE>
A:Cross-references: EMBL:L28003; NID:g2934883; PID:g2934884
A:Experimental source: strain Williams; hypocotyl

Query Match 58.0%; Score 47; DB 2; Length 326;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
||| | | | |
Db 61 CGFPLCHSC 69

RESULT 6

T47488 hypothetical protein F9K21.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47488
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, February 2000
A:Reference number: 224467
A:Accession: T47488
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-373 <COR>
A:Cross-references: EMBL:AL138657
A:Experimental source: Cultivar Columbia; BAC clone F9K21
C:Genetics:
A:Map position: 3
A:introns: 92/3; 328/2; 353/3
A:Note: F9K21.60

Query Match 58.0%; Score 47; DB 2; Length 373;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
||| | | | |
Db 329 CGYKLCYAC 337

RESULT 7

S25774 testis-specific protein Mst84Dc - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
C:Accession: S25774; C56565
R:Kuhn, R.; Kuhn, C.; Boersch, D.; Glaetzer, K.H.; Schaefer, U.; Schaefer, M.
Mech. Dev. 35, 143-151, 1991
A:Title: A cluster of four genes selectively expressed in the male germ line of Dros
A:Reference number: A56565; MUID:92102953; PMID:1684716
A:Accession: S25774
A:Molecule type: DNA
A:Residues: 1-55 <KUH>
A:Cross-references: EMBL:X67703; NID:g11072; PIDN:CAA47939.1; PID:g11075
A:Note: the authors translated the codon TGC for residue 55 as Thr
A:Note: sequence extracted from NCBI backbone (NCBIN:74217, NCBI:74222)
C:Genetics:
A:Gene: Mst84Dc
A:Cross-references: FlyBase:FBgn0004174
A:Map position: 3
C:Superfamily: fruit fly testis-specific protein
C:Keywords: spermatogenesis; tandem repeat

Query Match 56.8%; Score 46; DB 2; Length 55;
Best Local Similarity 58.3%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 GYCCGYKLCXHC 13
||| | | | |
Db 7 GSCCGYCCGFC 18

RESULT 8

P95210

conserved domain protein SP1806 [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: F95210
 R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: F95210
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-79 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75879.1; PID:q14973305; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP1806

Query Match 56.8%; Score 46; DB 2; Length 79;
 Best Local Similarity 63.6%; Pred. No. 8.4;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VGVCCGKLC H 11
 : || ||| ||
 Db 39 IKVCLGKRC H 49

RESULT 9
 T42321
 hypothetical protein 31.1 - phage SP1
 C:Species: phage SP1
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
 C:Accession: T42321
 R:Alonso, J.C.; Luder, G.; Stiege, A.C.; Chal, S.; Weise, F.; Trautner, T.A.
 Gene 204, 201-212, 1997
 A:Title: The complete nucleotide sequence and functional organization of Bacillus subtil
 A:Reference number: Z22137; MUID:98094274; PMID:9434185
 A:Accession: T42321
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-126 <ALO>
 A:Cross-references: EMBL:X97918; PIDN:CAA66528.1
 C:Superfamily: phage SP1 hypothetical protein 31.1

Query Match 55.6%; Score 45; DB 2; Length 126;
 Best Local Similarity 77.8%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VGVCCGKYL 9
 : ||||| ||
 Db 104 IGVCCGYL 112

RESULT 10
 T26913
 hypothetical protein Y45F10B.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T26913
 R:McMurray, A.
 submitted to the EMBL Data Library, January 1998
 A:Reference number: Z20286
 A:Accession: T26913
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-244 <WII>
 A:Cross-references: EMBL:AL021487; PIDN:CAA16351.1; GSPDB:GN00022; CESP:Y45F10B.8
 A:Experimental source: clone Y45F10B
 C:Genetics:
 A:Gene: CESP:Y45F10B.8
 A:Map position: 4

A:introns: 90/2; 194/1

Query Match 54.3%; Score 44; DB 2; Length 244;
 Best Local Similarity 55.6%; Pred. No. 35;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 CGYKLCXHC 13
 : ||: ||| |
 Db 195 CGHTTCHTC 203

RESULT 11
 T26912
 hypothetical protein Y45F10B.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T26912
 R:McMurray, A.
 submitted to the EMBL Data Library, January 1998
 A:Reference number: Z20286
 A:Accession: T26912
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-244 <WII>
 A:Cross-references: EMBL:AL021487; PIDN:CAA16350.1; GSPDB:GN00022; CESP:Y45F10B.9
 A:Experimental source: clone Y45F10B
 C:Genetics:
 A:Gene: CESP:Y45F10B.9
 A:Map position: 4
 A:introns: 90/2; 194/1

Query Match 54.3%; Score 44; DB 2; Length 244;
 Best Local Similarity 55.6%; Pred. No. 35;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 CGYKLCXHC 13
 : ||: ||| |
 Db 195 CGHTTCHTC 203

RESULT 12
 T24312
 hypothetical protein T01G5.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T24312
 R: Basham, V.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19873
 A:Accession: T24312
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-332 <WII>
 A:Cross-references: EMBL:Z81111; PIDN:CAB03269.1; GSPDB:GN00023; CESP:T01G5.7
 A:Experimental source: clone T01G5
 C:Genetics:
 A:Gene: CESP:T01G5.7
 A:Map position: 5
 A:introns: 208/1; 257/1

Query Match 54.3%; Score 44; DB 2; Length 332;
 Best Local Similarity 55.6%; Pred. No. 43;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 CGYKLCXHC 13
 : ||: ||| |
 Db 209 CGHTTCHTC 217

RESULT 13
 T10559
 hypothetical protein F25F4.10 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
C:Accession: T10559
R:Byevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft
submitted to the Protein Sequence Database, June 1999
A:Reference number: 216533
A:Accession: T10559
A:Molecule type: DNA
A:Residues: 1-514 <BEV>
A:Cross-references: EMBL:AL050399; GSPDB:GN00062; ATSP:F25E4.10
A:Experimental source: cultivar Columbia; BAC clone F25E4
C:Genetics:
A:Gene: ATSP:F25E4.10
A:Map position: 4
A:introns: 436/3; 457/3; 479/3

Query Match 54.3%; Score 44; DB 2; Length 514;
Best Local Similarity 60.0%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
||| :|| :|
DB 295 CCGFVLCFC 304

RESULT 14

g72867
hypothetical protein - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, AchNPV
A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
C:Accession: G72867
R:AYres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A:Reference number: A72850; MUID:94303173; PMID:8030224
A:Accession: G72867
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-261 <AYR>
A:Cross-references: GB:L22858; NID:9510708; PIDN:AAA6771.1; PID:9559210
C:Genetics:
A:Gene: Ac-IE-0

Query Match 53.1%; Score 43; DB 2; Length 261;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
||| :|| :|
DB 230 CCEVAICNAC 239

RESULT 15

IE-0 orf141 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C:Species: Bombyx mori nuclear polyhedrosis virus, BmsNPV
A:Variety: Isolate T3
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C:Accession: T41878
R:Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: 222020; MUID:99281911; PMID:10355780
A:Accession: T41878
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-261 <KAN>
A:Cross-references: EMBL:L33180; PIDN:AAC63807.1
A:Experimental source: isolate T3
C:Genetics:
A:Note: ie-0

Query Match 53.1%; Score 43; DB 2; Length 261;

Best Local Similarity 50.0%; Pred. No. 50;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 4 CCGYKLCXHC 13
||| :|| :|
DB 230 CCEVAICNAC 239

Search completed: December 2, 2002, 10:08:58
Job time : 14.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:04:20 ; Search time 7.5 Seconds
(without alignments)
71.892 Million cell updates/sec

Title: US-09-787-986A-2
Perfect score: 81
Sequence: 1 VGVCCGYKLCXHC 13

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	96.3	13	CXL4_CONMR	P58810 conus marino
2	74	91.4	61	CXL2_CONMR	P58808 conus marino
3	68	84.0	11	CXL1_CONMR	P58807 conus marino
4	50	61.7	12	CXL3_CONMR	P58809 conus marino
5	47	58.0	245	IEO_NPVOP	O10369 corygia pseu
6	46	56.8	55	MB4C_DROME	O01644 drosophila
7	45	55.6	465	RN15_HUMAN	O00635 homo sapien
8	43	53.1	261	IEO_NPVAC	P41710 autographa
9	43	53.1	554	ARD1_RAT	P36407 rattus norv
10	43	53.1	574	ARD1_HUMAN	P51795 homo sapien
11	43	53.1	746	CIC5_HUMAN	O9wvd4 mus musculu
12	43	53.1	746	CIC5_MOUSE	P51796 rattus norv
13	43	53.1	746	CIC5_RAT	P51796 rattus norv
14	42	51.9	172	PGPA_ECOLI	P18200 escherichia
15	42	51.9	207	CSF3_HUMAN	P09919 homo sapien
16	42	51.9	400	VG78_HSV1	O00167 icteralrid h
17	42	51.9	489	YAC4_SCHPO	O09818 schizosacch
18	42	51.9	587	NOT4_YEAST	P34909 saccharomyc
19	41	50.6	208	CSF3_MOUSE	P09920 mus musculu
20	41	50.6	1235	DDB1_HCMVA	P17147 human cytom
21	40	49.4	321	ABNA_ASPNG	P42256 aspergillus
22	40	49.4	469	AR11_MOUSE	O9z1k5 mus musculu
23	40	49.4	557	AR11_HUMAN	O9y4x5 homo sapien
24	40	49.4	560	PRK1_YEAST	P36002 saccharomyc
25	40	48.4	486	DIH9_HUMAN	O9nyc9 homo sapien
26	39.5	48.8	768	ITB8_RABIT	P26013 oryctolagus
27	39	48.1	218	YT44_CAEEL	O11072 caenorhabdi
28	39	48.1	221	TRM7_HUMAN	O9c029 homo sapien
29	39	48.1	257	RPEF_SALTY	P26396 salmoneilla
30	39	48.1	798	ITB1_XENLA	P12806 xenopus lae
31	39	48.1	973	RPEO_PAV	O9j7z2 pariacoto v
32	39	48.1	994	CIC1_MOUSE	O64347 mus musculu
33	39	48.1	994	CIC1_RAT	P35524 rattus norv

34	39	48.1	1191	1	DDB1_MCMVS	P30672 murine cyto
35	38.5	47.5	30	1	CXEX_CONCN	P58928 conus conso
36	38.5	47.5	1203	1	MGR5_RAT	P31424 rattus norv
37	38.5	47.5	1212	1	MGR5_HUMAN	P41594 homo sapien
38	38	46.9	72	1	YVAU_VACCC	P20530 vaccinia vi
39	38	46.9	155	1	NEU4_CATCO	P16229 catostomus
40	38	46.9	158	1	VE6_HPV45	P21735 human papil
41	38	46.9	194	1	CSF3_FELCA	O02708 felis silve
42	38	46.9	195	1	CSF3_BOVIN	P35833 bos taurus
43	38	46.9	224	1	IAPL_ASFB7	O65138 african swi
44	38	46.9	224	1	IAPL_ASFC3	O11451 african swi
45	38	46.9	224	1	IAPL_ASFC8	O12407 african swi

ALIGNMENTS

RESULT 1						
ID	CXL4_CONMR	STANDARD:	PRT:	13 AA.		
AC	P58810:					
DT	15-JUN-2002 (Rel. 41, Created)					
DT	15-JUN-2002 (Rel. 41, Last sequence update)					
DT	15-JUN-2002 (Rel. 41, Last annotation update)					
DE	Lambda/chi-conotoxin MrIB (Chi-MrIB).					
OS	Conus marmoreus (Marble cone).					
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;					
OC	Neogastropoda; Conoidea; Conidae; Conus.					
OX	NCBI_TaxID=42752;					
RN	[1]					
RP	SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.					
RC	TISSUE-Venom:					
RX	MEDLINE=21419681; PubMed=11528421;					
RA	Shaper I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,					
RA	Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,					
RA	Lewis R.J.;					
RT	"Two new classes of conopeptides inhibit the alpha-adrenoceptor and					
RT	norepinephrine transporter".					
RL	Nat. Neurosci. 4:902-907(2001).					
CC	- FUNCTION: Inhibits the neuronal noradrenergic transporter.					
CC	- SUBCELLULAR LOCATION: Secreted.					
CC	- TISSUE SPECIFICITY: Expressed by the venom duct.					
CC	- PTM: Exists in two forms, due to cis-trans isomerization at His-					
CC	11-Hyp-12.					
CC	- MASS SPECTROMETRY: MW=1393.52; METHOD=Electrospray.					
CC	- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.					
DR	PDB: 1IEO; 03-Apr-02.					
KW	Neurotoxin; Toxin; Hydroxylation; 3D-structure.					
FT	DISULFID 4 13					
FT	DISULFID 5 10					
FT	MOD.RES 12 12					
SQ	SEQUENCE 13 AA: 1382 MW: 277AAC376EAD2B58 CRC64;					
Query Match						
Best local similarity		96.3%;	Score 78;	DB 1;	Length 13;	
Matches 12;		Conservative	92.3%;	Pred. No. 2.3e-05;		
			0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1 VGVCCGYKLCXHC 13					
DB	1 VGVCCGYKLCXHC 13					
RESULT 2						
ID	CXL2_CONMR	STANDARD:	PRT:	61 AA.		
AC	P58808:					
DT	15-JUN-2002 (Rel. 41, Created)					
DT	15-JUN-2002 (Rel. 41, Last sequence update)					
DT	15-JUN-2002 (Rel. 41, Last annotation update)					
DE	Lambda-conotoxin CMrYIB precursor (Chi-conotoxin MrIA) (Chi-MrIA)					
DE	(Mr10a).					
OS	Conus marmoreus (Marble cone).					
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;					

OC Neogastropoda: Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 49-61, SYNTHESIS, AND MASS
RP SPECTROMETRY.
RC TISSUE-Venom duct, and Venom;
RX MEDLINE=20490660; PubMed=10900201;
RA McIntosh J.M., Corpuz G.O., Layer R.T., Garrett J.E., Wagstaff J.D.,
RA Bulaj G., Vazovkina A., Yoshikami D., Cruz L.J., Oliveira B.M.;
RT "Isolation and characterization of a novel conus peptide with apparent
RT antinociceptive activity.";
RL J. Biol. Chem. 275:32391-32397(2000).
RN [12]
RP SEQUENCE OF 49-61, AND MASS SPECTROMETRY.
RC TISSUE-Venom;
RX MEDLINE=20564325; PubMed=10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
RN [13]
RP SEQUENCE OF 49-61, SYNTHESIS, MASS SPECTROMETRY, AND STRUCTURE BY NMR.
RC TISSUE-Venom;
RX MEDLINE=21419681; PubMed=11528421;
RA Sharpe I.A., Gehmann J., Loughnan M.L., Thomas L., Adams D.A.,
RA Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,
RA Lewis R.J.;
RT "Two new classes of conopeptides inhibit the alpha1-adrenoceptor and
RT noradrenaline transporter.";
RL Nat. Neurosci. 4:902-907(2001).
CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- PTM: Exists in two forms, due to cis-trans isomerization at His-
CC 59-Hyp-60.
CC -1- MASS SPECTROMETRY: MW=1408.5; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KM Neurotoxin; Toxin; Hydroxylation; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 48
RX Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RX Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- MASS SPECTROMETRY: MW=1262.77; MW-ERR=0.07; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KM Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 3 12
FT MOD.RES 11 11 HYDROXYLATION.
SQ SEQUENCE 61 AA; 6499 MW; F4DE5B5A97EB8DA CRC64;
Query Match 91.4%; Score 74; DB 1; Length 61;
Best Local Similarity 91.7%; Pred. No. 0.00025;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GVCCGYKLCHXC 13
DB 50 GVCCGYKLCHPC 61
CXL1_CONMR STANDARD; PRT; 11 AA.
ID CXL1_CONMR
AC P58807;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CMYIA.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [11]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE-Venom;
RX MEDLINE=20564325; PubMed=10988292;

RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- MASS SPECTROMETRY: MW=1237.93; MW-ERR=0.21; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KM Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 2 11
FT MOD.RES 10 10 HYDROXYLATION.
SQ SEQUENCE 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;
Query Match 84.0%; Score 68; DB 1; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.00051;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 VCCGYKLCHXC 13
DB 1 VCCGYKLCHPC 11
CXL3_CONMR STANDARD; PRT; 12 AA.
ID CXL3_CONMR
AC P58809;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CMYX.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [11]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE-Venom;
RX MEDLINE=20564325; PubMed=10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- MASS SPECTROMETRY: MW=1262.77; MW-ERR=0.07; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KM Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 3 12
FT MOD.RES 11 11 HYDROXYLATION.
SQ SEQUENCE 12 AA; 1251 MW; 277AAE242D5A2C68 CRC64;
Query Match 61.7%; Score 50; DB 1; Length 12;
Best Local Similarity 50.0%; Pred. No. 0.18;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 GVCCGYKLCHXC 13
DB 1 GICCGVSPCTPC 12
IEO_NPVOP STANDARD; PRT; 245 AA.
ID IEO_NPVOP
AC O10369;
DT 01-NOV-1997 (Rel. 35, Created)


```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Immediate-early protein IE-0.
GN IE-0.
OS Orygia pseudotsugata multicaudid polyhedrosis virus (OPMPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolydiovirus.
OX NCBI_TaxID=164623;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohrmann G.F.;
RT "The sequence of the Orygia pseudotsugata multicaudid nuclear
RT polyhedrosis virus genome.";
RL Virology 229:381-398(1997).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U75930; AAC59137.1; -
DR InterPro: IPR001841; Znf_ring.
DR PROSITE: PS00518; ZF_RING_1; FALSE-NEG.
DR PROSITE: PS00089; ZF_RING_2; 1.
KW Early protein; Zinc-finger.
FT 2N.FING 195 240 RING-TYPE.
SO SEQUENCE 245 AA; 27117 MW; 357F43B15F7B1029 CRC64;

Oy 4 CCGYKLCXHC 13
Db 213 CCGYSLCYAC 222

Query Match 58.0%; Score 47; DB 1; Length 245;
Best Local Similarity 60.0%; Pred. No. 4.1;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 6
M84C_DROME STANDARD: PRT: 55 AA.
AC Q01644; OSVINO;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Male specific sperm protein Mst84C.
GN Mst84C OR CG17945.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92102953; PubMed=1684716;
RA Kuhn R., Kuhn C., Boersch D., Glaetzer K.H., Schaefer U.,
RA Schaefer M.;
RT "A cluster of four genes selectively expressed in the male germ line
RT of Drosophila melanogaster.";
RL Mech. Dev. 35:143-151(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

```

```

RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abdl J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova K.C., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Meltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissensbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.
CC -1- DOMAIN: THIS PROTEIN IS MOSTLY COMPOSED OF REPETITIVE C-G-P
CC MOTIFS.
CC -----
CC -1- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67703; AAF47939.1; -
DR EMBL: AE003672; AAF54025.1; -
DR FlyBase: FBgn004174; Mst84C.
KW Spermatogenesis; Repeat; Multigene family.
SO SEQUENCE 55 AA; 5225 MW; 95A12F3AEC88BD6C CRC64;

Query Match 56.8%; Score 46; DB 1; Length 55;
Best Local Similarity 58.3%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 2 GVCYKLCXHC 13
Db 7 GSCGYCCGCP 18

RESULT 7
RN15_HUMAN STANDARD: PRT: 465 AA.
AC Q00635;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RING finger protein 15 (Zinc-finger protein Rofet) (Tripartite motif-
DE containing protein 38).

```

```

GN TRIM38 OR RNF15 OR RORET.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97294057; PubMed=9149941;
RA Ruddy D.A., Krommal G.S., Lee V.K., Muntier G.A., Quintana L.,
RA Domingo R., Jr., Meyer N.C., Irrinki A., McCelland E.E., Fullan A.,
RA Mapa F.A., Moore T., Thomas W., Loeb D.B., Harmon C., Tsuchihashi Z.,
RA Wolf R.K., Schatzman R.C., Feder J.N.;
RT "A 1.1-Mb transcribed map of the hereditary hemochromatosis locus.";
RL Genome Res. 7:441-456(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -1- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC -1- SIMILARITY: CONTAINS 1 SPRY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U91328; AAB82084.1; -
DR EMBL: U90547; AAB53425.1; -
DR EMBL: BC026930; AAB26930.1; -
DR Genew; HGNC:10059; TRIM38.
DR InterPro: IPR001870; Gamma_carboxylase.
DR InterPro: IPR003877; SPRY.
DR InterPro: IPR003878; SPRY_domain.
DR InterPro: IPR000315; Znf_Box.
DR InterPro: IPR001841; Znf_finger.
DR Pfam: PF00622; SPRY.1.
DR Pfam: PF00643; Zf-B_Box.1.
DR Pfam: PF00097; Zf-C3HC4.1.
DR PRINTS: PR01406; BBOXZNFINGER.
DR SMART: SM00336; BBOX.1.
DR SMART: SM00184; RING.1.
DR SMART: SM00449; SPRY.1.
DR PROSITE: PS00119; ZF_BOX.1.
DR PROSITE: PS00518; ZF_RING.1.
DR PROSITE: PS00089; ZF_RING.2.1.
KW Zinc-finger; Polymorphism.
FT ZN_FING 16 63 RING-TYPE.
FT ZN_FING 88 129 B_BOX-TYPE.
FT DOMAIN 344 464 SPRY.
FT VARIANT 421 421 G->R (IN DBSNP:8469041).
FT CONFLICT 23 23 /FTID=VAR_013513.
FT SEQUENCE 465 AA; 53416 MM; 30A884051A2DA058 CRC64;
SQ

```

```

Query Match 55.6%; Score 45; DB 1; Length 465;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 1 VGVGGYKLCXHC 13
DB 27 VSINGCHSYCLC 39

```

```

RESULT 8
ID IEQ_NPVAC STANDARD: PRT: 261 AA.
AC P41710.

```

```

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Immediate-early protein IE-0.
OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolydnavirus.
OX NCBI_TaxID=46015;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C6;
RX MEDLINE=94303173; PubMed=8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
RT polyhedrosis virus.";
RL Virology 202:586-605(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L2858; AAB6771.1; -
DR InterPro: IPR001841; Znf_finger.
DR PROSITE: PS00518; ZF_RING.1; FALSE-NEG.
DR PROSITE: PS50089; ZF_RING.2.1.
KW Early protein; Zinc-finger.
FT ZN_FING 212 257 RING-TYPE.
FT SEQUENCE 261 AA; 30109 MM; 7721E0C558EC2CBE CRC64;
SQ

```

```

Query Match 53.1%; Score 43; DB 1; Length 261;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 4 CGGYKLCXHC 13
DB 230 CCEVATCNC 239

```

```

RESULT 9
ID ARD1_RAT STANDARD: PRT: 554 AA.
AC P36407.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP-binding protein ARD-1 (Fragment).
GN ARD1 OR ARD1 OR ARD-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93232038; PubMed=8473324;
RA Mishima K., Tsuchiya M., Nightingale M.S., Moss J., Vaughan M.;
RT "ARD 1, a 64-kDa guanine nucleotide-binding protein with a carboxyl-
RT terminal ADP-ribosylation factor domain.";
RL J. Biol. Chem. 268:8801-8807(1993).
CC -1- FUNCTION: NOT KNOWN. THE C-TERMINUS CAN ACT AS AN ALLOSTERIC
CC ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ARF FAMILY
CC OF GTP-BINDING PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -1- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```


DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Chloride channel protein 5 (CLC-5).
 GN CLCN5 OR CLCK2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID:9606.
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE-Kidney;
 RC MEDLINE=96121370; PubMed=8575751;
 RX Fisher S.E., van Bakel I., Lloyd S.E., Pearce S.H.S.,
 RA Thakker R.V., Craig I.W.,
 RA Thakker R.V., Craig I.W.;
 RT "Cloning and characterization of CLCN5, the human kidney chloride
 channel gene implicated in Dent disease (an X-linked hereditary
 nephrolithiasis).";
 RL Hum. Mol. Genet. 3:598-606(1995).
 [2]
 RN SEQUENCE OF 487-746 FROM N.A.
 RP TISSUE-Kidney;
 RC MEDLINE=95179126; PubMed=7874126;
 RX Fisher S., Black G.C.M., Lloyd S.E., Hatchwell E., Wrong O.,
 RA Thakker R.V., Craig I.W.;
 RT "Isolation and partial characterization of a chloride channel gene
 which is expressed in kidney and is a candidate for Dent's disease
 (an X-linked hereditary nephrolithiasis).";
 RL Hum. Mol. Genet. 3:2053-2059(1994).
 [3]
 RN TISSUE SPECIFICITY.
 RP TISSUE=Vascular smooth muscle, and Aortic endothelium;
 RC MEDLINE=99222497; PubMed=10198195;
 RX Lamb F.S., Clayton G.H., Liu B.-X., Smith R.L., Barna T.J.,
 RA Schutte B.C.;
 RT "Expression of CLCN voltage-gated chloride channel genes in human
 blood vessels";
 RL J. Mol. Cell. Cardiol. 31:657-666(1999).
 [4]
 RN VARIANTS NPHL ARG-200; LEU-244; GLU-506 AND PRO-520.
 RP MEDLINE=96158876; PubMed=8559248;
 RX Lloyd S.E., Pearce S.H.S., Fisher S.E., Steinmeyer K., Schwappach B.,
 RA Schelman S.J., Harding B., Bollno A., Devoto M., Goody P.,
 RA Rigden S.P.A., Wrong O., Jentsch T.J., Craig I.W., Thakker R.V.;
 RT "A common molecular basis for three inherited kidney stone diseases.";
 RL Nature 379:445-449(1996).
 [5]
 RN VARIANTS NPHL.
 RP MEDLINE=97402204; PubMed=9259268;
 RX Lloyd S.E., Guenther W., Pearce S.H.S., Thomson A., Bianchi M.L.,
 RA Bosio M., Craig I.W., Fisher S.E., Schelman S.J., Wrong O.,
 RA Jentsch T.J., Thakker R.V.;
 RT "Characterisation of renal chloride channel, CLCN5, mutations in
 hypercalcaemic nephrolithiasis (kidney stones) disorders.";
 RL Hum. Mol. Genet. 6:1233-1239(1997).
 [6]
 RN FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
 SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
 MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
 TRANSEPITHELIAL TRANSPORT. MAY PLAY AN IMPORTANT ROLE IN RENAL
 TUBULAR FUNCTION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: KIDNEY. MODERATELY EXPRESSED IN AORTIC
 VASCULAR SMOOTH MUSCLE AND ENDOTHELIAL CELLS, AND AT A SLIGHTLY
 CC -1- HIGHER LEVEL IN THE CORONARY VASCULAR SMOOTH MUSCLE.
 CC -1- DISEASE: DEFECTS IN CLCN5 ARE THE CAUSE OF FOUR DISORDERS OF
 HEREDITARY HYPERCALCAEMIC NEPHROLITHIASIS (KIDNEY STONES), THAT
 CC HAVE BEEN REFERRED TO AS DENT'S DISEASE (DD), X-LINKED RECESSIVE
 NEPHROLITHIASIS (XRN), X-LINKED RECESSIVE HYPOPHOSPHATAMIC
 CC RICKETS (XLHR) AND IDIOPATHIC LOW MOLECULAR WEIGHT PROTEINURIA OF
 CC JAPANESE CHILDREN (JILP). ALL FOUR DISEASES REPRESENT RENAL
 CC TUBULAR DISORDERS THEY ARE CHARACTERIZED BY LOW MOLECULAR WEIGHT
 CC PROTEINURIA, HYPERCALCAEMIA, NEPHROCALCINOSIS, NEPHROLITHIASIS
 CC (KIDNEY STONES) AND RENAL FAILURE. DD IS A FORM OF FANCONI
 CC SYNDROME (ALSO KNOWN AS X-LINKED RECESSIVE NEPHROLITHIASIS TYPE 2
 CC (NPHL2)).

CC	-1- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.
CC	-1- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC	-----
CC	EMBL; X91906; CAA63000.1; -.
DR	EMBL; X81836; CAA57430.1; -.
DR	Genew; HGNC:2023; CLCN5.
DR	MIM; 300008; -.
DR	MIM; 300009; -.
DR	MIM; 310468; -.
DR	InterPro; IPR000644; CBS_domain.
DR	InterPro; IPR001807; Cl-channel_volt.
DR	Pfam; PF00571; CBS; 2.
DR	Pfam; PF00654; voltage_CLC; 1.
DR	PRINTS; PR00762; CLCHANNEL.
DR	SMART; SM00116; CBS; 2.
KM	Ionic channel; Ion transport; Voltage-gated channel; Transmembrane; .
KW	CBS domain; Repeat; Disease mutation.
FT	TRANSMEM 55 POTENTIAL.
FT	TRANSMEM 133 153 POTENTIAL.
FT	TRANSMEM 189 209 POTENTIAL.
FT	TRANSMEM 246 266 POTENTIAL.
FT	TRANSMEM 317 337 POTENTIAL.
FT	TRANSMEM 352 372 POTENTIAL.
FT	TRANSMEM 428 448 POTENTIAL.
FT	TRANSMEM 453 473 POTENTIAL.
FT	TRANSMEM 494 514 POTENTIAL.
FT	TRANSMEM 521 541 POTENTIAL.
FT	DOMAIN 585 645 CBS 1.
FT	DOMAIN 682 733 CBS 2.
FT	VARIANT 30 R -> RH (IN DD).
FT	VARIANT 57 /FTId=VAR_001615.
FT	VARIANT 200 L -> R (IN NPHI2).
FT	VARIANT 244 S -> L (IN NPHI3).
FT	VARIANT 280 /FTId=VAR_001618.
FT	VARIANT 506 R -> P (IN JIIP).
FT	VARIANT 512 /FTId=VAR_001619.
FT	VARIANT 520 G -> E (IN NPHI1).
FT	VARIANT 527 /FTId=VAR_001620.
FT	VARIANT 527 G -> R (IN DD); ABOLISHES THE CHLORIDE CURRENTS).
FT	VARIANT 527 /FTId=VAR_001621.
FT	VARIANT 527 /FTId=VAR_001622.
FT	VARIANT 527 E -> D (IN DD); ABOLISHES THE CHLORIDE CURRENTS AND TOTAL LOSS OF FUNCTION).
FT	VARIANT 527 /FTId=VAR_001623.
FT	VARIANT 527 /FTId=VAR_001624.
FT	VARIANT 527 /FTId=VAR_001625.
FT	VARIANT 527 /FTId=VAR_001626.
FT	VARIANT 527 /FTId=VAR_001627.
FT	VARIANT 527 /FTId=VAR_001628.
FT	VARIANT 527 /FTId=VAR_001629.
FT	VARIANT 527 /FTId=VAR_001630.
FT	VARIANT 527 /FTId=VAR_001631.
FT	VARIANT 527 /FTId=VAR_001632.
FT	VARIANT 527 /FTId=VAR_001633.
FT	VARIANT 527 /FTId=VAR_001634.
FT	VARIANT 527 /FTId=VAR_001635.
FT	VARIANT 527 /FTId=VAR_001636.
FT	VARIANT 527 /FTId=VAR_001637.
FT	VARIANT 527 /FTId=VAR_001638.
FT	VARIANT 527 /FTId=VAR_001639.
FT	VARIANT 527 /FTId=VAR_001640.
FT	VARIANT 527 /FTId=VAR_001641.
FT	VARIANT 527 /FTId=VAR_001642.
FT	VARIANT 527 /FTId=VAR_001643.
FT	VARIANT 527 /FTId=VAR_001644.
FT	VARIANT 527 /FTId=VAR_001645.
FT	VARIANT 527 /FTId=VAR_001646.
FT	VARIANT 527 /FTId=VAR_001647.
FT	VARIANT 527 /FTId=VAR_001648.
FT	VARIANT 527 /FTId=VAR_001649.
FT	VARIANT 527 /FTId=VAR_001650.
FT	VARIANT 527 /FTId=VAR_001651.
FT	VARIANT 527 /FTId=VAR_001652.
FT	VARIANT 527 /FTId=VAR_001653.
FT	VARIANT 527 /FTId=VAR_001654.
FT	VARIANT 527 /FTId=VAR_001655.
FT	VARIANT 527 /FTId=VAR_001656.
FT	VARIANT 527 /FTId=VAR_001657.
FT	VARIANT 527 /FTId=VAR_001658.
FT	VARIANT 527 /FTId=VAR_001659.
FT	VARIANT 527 /FTId=VAR_001660.
FT	VARIANT 527 /FTId=VAR_001661.
FT	VARIANT 527 /FTId=VAR_001662.
FT	VARIANT 527 /FTId=VAR_001663.
FT	VARIANT 527 /FTId=VAR_001664.
FT	VARIANT 527 /FTId=VAR_001665.
FT	VARIANT 527 /FTId=VAR_001666.
FT	VARIANT 527 /FTId=VAR_001667.
FT	VARIANT 527 /FTId=VAR_001668.
FT	VARIANT 527 /FTId=VAR_001669.
FT	VARIANT 527 /FTId=VAR_001670.
FT	VARIANT 527 /FTId=VAR_001671.
FT	VARIANT 527 /FTId=VAR_001672.
FT	VARIANT 527 /FTId=VAR_001673.
FT	VARIANT 527 /FTId=VAR_001674.
FT	VARIANT 527 /FTId=VAR_001675.
FT	VARIANT 527 /FTId=VAR_001676.
FT	VARIANT 527 /FTId=VAR_001677.
FT	VARIANT 527 /FTId=VAR_001678.
FT	VARIANT 527 /FTId=VAR_001679.
FT	VARIANT 527 /FTId=VAR_001680.
FT	VARIANT 527 /FTId=VAR_001681.
FT	VARIANT 527 /FTId=VAR_001682.
FT	VARIANT 527 /FTId=VAR_001683.
FT	VARIANT 527 /FTId=VAR_001684.
FT	VARIANT 527 /FTId=VAR_001685.
FT	VARIANT 527 /FTId=VAR_001686.
FT	VARIANT 527 /FTId=VAR_001687.
FT	VARIANT 527 /FTId=VAR_001688.
FT	VARIANT 527 /FTId=VAR_001689.
FT	VARIANT 527 /FTId=VAR_001690.
FT	VARIANT 527 /FTId=VAR_001691.
FT	VARIANT 527 /FTId=VAR_001692.
FT	VARIANT 527 /FTId=VAR_001693.
FT	VARIANT 527 /FTId=VAR_001694.
FT	VARIANT 527 /FTId=VAR_001695.
FT	VARIANT 527 /FTId=VAR_001696.
FT	VARIANT 527 /FTId=VAR_001697.
FT	VARIANT 527 /FTId=VAR_001698.
FT	VARIANT 527 /FTId=VAR_001699.
FT	VARIANT 527 /FTId=VAR_001700.
FT	

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chloride channel protein 5 (CLC-5).
GN CLC5 OR CLC5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=99303559; PubMed=10373326;
RA Tanaka K., Fisher S.E., Craig I.W.;
RT "Characterization of novel promoter and enhancer elements of the mouse
homologue of the Dent disease gene, CLCNS, implicated in X-linked
hereditary nephrolithiasis.";
RL Genomics 58:281-292(1999).
CC -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
TUBULAR FUNCTION. MAY PLAY AN IMPORTANT ROLE IN RENAL
CELLULAR FUNCTION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: KIDNEY-SPECIFIC.
CC -!- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CBS DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC
DR EMBL, AF134117; AAD28473.1; -.
DR MGD; MGI:99486; CLCNS.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR001807; Cl-channel_volt.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF00654; voltage_CLC; 1.
DR PRINTS: PR00762; CLCHANNEL.
DR SMART: SM00116; CBS; 2.
KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
KW CBS domain; Repeat.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
FT TRANSMEM 521 541 POTENTIAL.
FT TRANSMEM 585 614 CBS 1.
FT DOMAIN 682 733 CBS 2.
FT DOMAIN 746 AA; 83100 MM; DBP3AE4FC331A08 CRC64;
SQ SEQUENCE

Query Match 53.1%; Score 43; DB 1; Length 746;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VGVCCGYKLCH 11
| | | | | | | | | |
DB 217 VACCCGNILCH 227

RESULT 13
CLC5_RAT
ID CLC5_RAT STANDARD; PRT; 746 AA.
AC P51796; P70642;
DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chloride channel protein 5 (CLC-5).
GN CLCNS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96125100; PubMed=8537381;
RA Steinmeyer K., Schwappach B., Bens M., Vandewalle A., Jentsch T.J.;
RT "Cloning and functional expression of rat CLC-5, a chloride channel
related to kidney disease.";
RL J. Biol. Chem. 270:31172-31177(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96215316; PubMed=8626585;
RA Sakamoto H., Kawasaki M., Uchida S., Sasaki S., Marumo F.;
RT "Identification of a new outwardly rectifying Cl⁻ channel that
belongs to a subfamily of the CLC Cl⁻ channels.";
RL J. Biol. Chem. 271:10210-10216(1996).
CC -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
TUBULAR FUNCTION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: KIDNEY-SPECIFIC.
CC -!- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CBS DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC
DR EMBL, Z56277; CAA91216.1; -.
DR EMBL, D50497; BAA09091.1; -.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR001807; Cl-channel_volt.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF00654; voltage_CLC; 1.
DR PRINTS: PR00762; CLCHANNEL.
DR SMART: SM00116; CBS; 2.
KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
KW CBS domain; Repeat.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
FT TRANSMEM 521 541 POTENTIAL.
FT TRANSMEM 585 614 CBS 1.
FT DOMAIN 682 733 CBS 2.
FT DOMAIN 746 AA; 83067 MM; 5F17D45F397003CE CRC64;
SQ SEQUENCE

Query Match 53.1%; Score 43; DB 1; Length 746;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VGVCCGYKLCH 11
| | | | | | | | | |
FT CONFLICT 315 315 H -> Y (IN REF. 2).
FT CONFLICT 315 315

Db 217 VACCIGNILCH 227

RESULT 14

PGPA_ECOLI STANDARD: PRT: 172 AA.

AC P18200: P77321: (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 35, Last sequence update)

DT 01-NOV-1987 (Rel. 40, Last annotation update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Phosphatidyglycerophosphatase A (EC 3.1.3.27).

GN PGPA OR B0418.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OC NCBI_TaxID=562;

OX NCBI_TaxID=562;

RN [1]

RN SEQUENCE FROM N.A.

RX MEDLINE=89033892; PubMed=2846510;

RA Icho T.;

RT "Membrane-bound phosphatases in Escherichia coli: sequence of the pga gene.";

RL J. Bacteriol. 170:5110-5116(1988).

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F., Riegler J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RA "The complete genome sequence of Escherichia coli K-12.";

RT Science 277:1453-1474(1997).

RN [3]

RN SEQUENCE FROM N.A.

RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K., Duncan M., Federspiel N., Hymen R., Kalman S., Komp C., Kurl O., Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;

RA Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.

RN [4]

RN SEQUENCE FROM N.A.

RC STRAIN-K12 / W3110;

RA Iida A., Hayashi M., Fujio T., Teshiba S.;

RL Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: ONE OF THE THREE PHOSPHOLIPID PHOSPHATASES, SPECIFICALLY HYDROLYZES PHOSPHATIDYGLYCEROPHOSPHATE.

CC -1- CATALYTIC ACTIVITY: Phosphatidylglycerophosphate + H(2)O = phosphatidylglycerol + phosphate.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN BY A FRAMESHIFT IN POSITION 42. IN ADDITION THE AUTHOR OF REF.1 HAS TRANSLATED THE WRONG DNA STRAND THUS PRODUCING AN ORF WHICH HAS NOTHING TO DO WITH THE ONE SHOWN HERE.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M23546; AAA24325.1; ALT_SEQ.

DR EMBL: AE000148; AAC73521.1; -.

DR EMBL: U82664; AAB40174.1; -.

DR EMBL: D17333; BAA21779.1; -.

DR PIR: A30192; PAECGA.

DR Ecogene: EGI0704; pgaA.

KW Hydrolase; Phospholipid degradation; Transmembrane; Inner membrane; Complete proteome.

FT TRANSMEM 32 52 POTENTIAL.

FT TRANSMEM 54 74 POTENTIAL.

FT TRANSMEM 142 162 POTENTIAL.

SEQ SEQUENCE 172 AA; 19418 MW; 9DA1C817CA36C8B9 CRC64;

Query Match 51.9%; Score 42; DB 1; Length 172;

Best Local Similarity 54.5%; Pred. No. 16;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VGVCCGYKRLCH 11

DB 66 LGICIGVYLCH 76

RESULT 15

CSF3_HUMAN STANDARD: PRT: 207 AA.

ID CSF3_HUMAN

AC P09919:

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Granulocyte colony-stimulating factor precursor (G-CSF) (Pluripolein)

DE (Filgrastim) (Lenograstim).

GN CSF3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RN SEQUENCE FROM N.A.

RX MEDLINE=86118679; PubMed=3484805;

RA Nagata S., Tsuchiya M., Asano S., Kaziro Y., Yamazaki T., Yamamoto O., Hirata Y., Kubota N., Oheda M., Nomura H., Yamazaki T.;

RA "Molecular cloning and expression of cDNA for human granulocyte colony-stimulating factor.";

RT Nature 319:415-418(1986).

RN [2]

RN SEQUENCE FROM N.A.

RX MEDLINE=86220137; PubMed=2423327;

RA Nagata S., Tsuchiya M., Asano S., Yamamoto O., Hirata Y., Kubota N., Oheda M., Nomura H., Yamazaki T.;

RA "The chromosomal gene structure and two mRNAs for human granulocyte colony-stimulating factor.";

RT EMBO J. 5:575-581(1986).

RN [3]

RN SEQUENCE FROM N.A.

RX MEDLINE=87196936; PubMed=3494801;

RA Devlin J.J., Devlin P.E., Myambo K., Lilly M.B., Rado T.A., Warren M.K.;

RA "Expression of granulocyte colony-stimulating factor by human cell lines.";

RT J. Leukoc. Biol. 41:302-306(1987).

RL [4]

RL SEQUENCE FROM N.A., AND VARIANTS MET-157 AND THR-174.

RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q., Nickerson D.A.;

RA Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.

RN [5]

RN SEQUENCE OF 19-207 FROM N.A.

RX MEDLINE=86151684; PubMed=2420009;

RA Souza L.M., Boone T.C., Gabriello J., Lai P.H., Zsebo K.M., Murdock D.C., Chazin V.R., Brzusewski J., Lu H., Chen K.K., Barand J., Platzer E., Moore M.A.S., Mertelsmann R., Welte K.;

RA "Recombinant human granulocyte colony-stimulating factor: effects on normal and leukemic myeloid cells.";

RT Science 232:61-66(1986).

RL [6]

RL CARBOHYDRATE-LINKAGE SITE.

RP MEDLINE=93293942; PubMed=7685769;

RA Clogston C.L., Hu S., Boone T.C., Lu H.S.;

RA "Glycosidase digestion, electrophoresis and chromatographic analysis of recombinant human granulocyte colony-stimulating factor glycoforms produced in Chinese hamster ovary cells.";

RT J. Chromatogr. A 637:55-62(1993).

RL [7]

RL STRUCTURE BY NMR.

RX MEDLINE-93106200; PubMed-1281794;
RA Zink T., Ross A., Ambrosius D., Rudolph R., Holak T.A.;
RT "Secondary structure of human granulocyte colony-stimulating factor
RT derived from NMR spectroscopy.";
RL FEBS Lett. 314:435-439(1992).
RN [8]
RP STRUCTURE BY NMR.
RX MEDLINE-94304859; PubMed-7518249;
RA Zink T., Ross A., Luers K., Cieslar C., Rudolph R., Holak T.A.;
RT "Structure and dynamics of the human granulocyte colony-stimulating
RT factor determined by NMR spectroscopy. Loop mobility in a four-helix-
RT bundle protein.";
RL Biochemistry 33:8453-8463(1994).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE-93281718; PubMed-7685117;
RA Hill C.P., Oselund T.D., Eisenberg D.;
RT "The structure of granulocyte-colony-stimulating factor and its
RT relationship to other growth factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:5167-5171(1993).
CC -1- FUNCTION: GRANULOCYTE/MACROPHAGE COLONY-STIMULATING FACTORS ARE
CC CYTOKINES THAT ACT IN HEMATOPOIESIS BY CONTROLLING THE PRODUCTION,
CC DIFFERENTIATION, AND FUNCTION OF 2 RELATED WHITE CELL POPULATIONS
CC OF THE BLOOD, THE GRANULOCYTES AND THE MONOCYTES-MACROPHAGES. THIS
CC CSF INDUCES GRANULOCYTES.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PTM: O-Glycan consists of gal-galnac disaccharide which can be
CC modified with up to two sialic acid residues (done in
CC recombinantly expressed g-CSF from CHO cells).
CC PHARMACEUTICAL: Available under the names Neupogen or Granulokine
CC (Amgen/Roche) and Granocyte (Rhône-Poulenc). Used to treat
CC neutropenia (a disorder characterized by an extremely low number
CC of neutrophils in blood).
CC -1- SIMILARITY: BELONGS TO THE IL-6 SUPERFAMILY.
CC -1- CAUTION: REF.4 MISQUOTES THE GENE NAME AS "CSF1".
CC -1- DATABASE: NAME=Neupogen/Granulokine;
CC NOTE=Clinical information on Neupogen/Granulokine:
CC WWW="http://www.neupogen.ch/monograph/frame0.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X03438; CAA27168.1; -;
DR EMBL: M13008; AAA03056.1; -;
DR EMBL: X03656; CAA27291.1; -;
DR EMBL: X03655; CAA27290.1; -;
DR EMBL: AF388025; AAK62469.1; -;
DR EMBL: M17706; AAA35882.1; -;
DR PIR: A24573; A24573.
DR PIR: A25093; A25093.
DR PDB: 1RHG; 31-JAN-94.
DR PDB: 1GNC; 31-JUL-94.
DR Genew: HGNC:2438; CSF3.
DR MIM: 138970; -;
DR InterPro: IPR003629; GCSF_MGF.
DR InterPro: IPR003573; IL6_MGF_GCSF.
DR Pfam: PF00489; IL6; 1.
DR PRINTS: PR00433; IL6GCSFMGF.
DR PRODOM: PD008388; GCSF_MGF; 1.
DR SMART: SM00126; IL6; 1.
DR DR PROSITE: PS00254; INTERLEUKIN_6; 1.
KW Cytokine; Growth factor; Glycoprotein; Signal; Alternative splicing;
KW Polymorphism; Pharmaceutical; 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 207 GRANULOCYTE COLONY-STIMULATING FACTOR.
FT DISULFD 69 75
FT DISULFD 97 107

FT CARBOHYD 166 166 O-LINKED (GALNAC. .) (BY SIMILARITY).
FT VARSPPLIC 66 68 MISSING (IN SHORT ISOFORM).
FT VARIANT 157 157 L->M.
FT /FTID=VAR_013073.
FT VARIANT 174 174 A->T.
FT /FTID=VAR_013074.
FT HELIX 41 65
FT HELIX 69 71
FT HELIX 77 86
FT TURN 87 88
FT HELIX 105 124
FT TURN 125 127
FT TURN 130 132
FT HELIX 133 156
FT TURN 157 158
FT HELIX 176 203
FT TURN 204 204
SQ SEQUENCE 207 AA; 22293 MW; 421F635ECC776996 CRC64;

Query Match 51.9%; Score 42; DB 1; Length 207;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 VGVCGGYKLCH 11
| | | | |
Db 66 VSECATYKILCH 76

Search completed: December 2, 2002, 10:07:28
Job time : 8.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 10:05:01 ; Search time 24.5 Seconds
(without alignments)
109.331 Million cell updates/sec

Title: US-09-787-986a-2
Perfect score: 81
Sequence: 1 VGVCCGYKLCXHC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	70.4	258	12	036453
2	56	69.1	151	12	010614
3	56	69.1	285	12	039H38
4	56	69.1	285	12	08V5X4
5	54	66.7	244	12	091B09
6	51	63.0	234	12	080L66
7	49	60.5	289	12	091B87
8	48	59.3	64	5	09BPE9
9	48	59.3	514	5	095Q73
10	47	58.0	112	12	081680
11	47	58.0	322	10	039896
12	47	58.0	326	10	039895
13	47	58.0	364	12	081547
14	47	58.0	373	10	09M1P9
15	46	58.0	446	4	096BQ3
16	46	56.8	69	2	09R6Z5

17	46	56.8	79	16	097P38	097P38 streptococc
18	46	56.8	121	11	09CZW7	09CZW7 mus musculu
19	46	56.8	372	11	099ND4	099ND4 rattus norv
20	46	56.8	538	4	0961E1	0961E1 homo sapien
21	46	56.8	547	11	P97472	P97472 mus musculu
22	46	56.8	547	11	091XC0	091XC0 mus musculu
23	45	55.6	105	2	085226	085226 pseudomonas
24	45	55.6	126	9	048481	048481 bacterioph
25	45	55.6	220	5	08SUS9	08SUS9 encephalito
26	45	55.6	721	12	091IL5	091IL5 white spot
27	45	55.6	1009	12	08VAC3	08VAC3 white spot
28	44	54.3	243	12	091GD2	091GD2 epiphyas po
29	44	54.3	244	5	062463	062463 caenorhabdi
30	44	54.3	244	5	062464	062464 caenorhabdi
31	44	54.3	332	5	018012	018012 caenorhabdi
32	44	54.3	485	4	096PF7	096PF7 homo sapien
33	44	54.3	485	4	08WZ70	08WZ70 homo sapien
34	44	54.3	514	10	09LD86	09LD86 arabidopsis
35	44	54.3	1441	10	091K63	091K63 arabidopsis
36	44	54.3	1466	10	08VZ24	08VZ24 arabidopsis
37	43	53.1	123	10	09S703	09S703 zea mays (m
38	43	53.1	261	12	092493	092493 bombyx mori
39	43	53.1	272	5	09U1S9	09U1S9 caenorhabdi
40	43	53.1	310	11	09WUG1	09WUG1 cavia porce
41	43	53.1	338	5	09USC2	09USC2 caenorhabdi
42	43	53.1	397	2	086938	086938 streptomyce
43	43	53.1	557	5	009654	009654 caenorhabdi
44	43	53.1	598	16	099Z03	099Z03 streptococc
45	43	53.1	746	6	09TUT3	09TUT3 oryctolagus

ALIGNMENTS

RESULT 1
ID 036453 PRELIMINARY: PRT; 258 AA.
AC 036453: 09YMW3;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Immediate early 0 protein (Immediate early transactivator 0).
GN IE-0.
OS Lymantria dispar multicapsid nuclear polyhedrosis virus (LdMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10449;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97445058: PubMed-9300047;
RA Pearson M.N., Rohrmann G.F.;
RT "Splicing is required for transactivation by the immediate early gene
1 of the Lymantria dispar multinucleocapsid nuclear polyhedrosis
virus.";
RL Virology 235:153-165(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99124785: PubMed-9887315;
RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA Slavicek J.M., Rohrmann G.F.;
RT "Sequence and analysis of the genome of a baculovirus pathogenic for
Lymantria dispar.";
RL Virology 253:17-34(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA Slavicek J., Rohrmann G.F.;
RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX Kuzio J.;
RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF006656; AAC58234.1; -.

```

DR EMBL: AF081810; AAC70206.1; -.
DR InterPro: IPR001841; Znf_ring.
DR SMART: SM00184; RING: 1.
SQ SEQUENCE 258 AA; 29395 MW; AC564CDF92282BAD CRC64;

Query Match          70.4%; Score 57; DB 12; Length 258;
Best Local Similarity 63.6%; Pred. No. 0.095;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 CCGYKLCXHC 13
   |||||:|:|
Db 208 VCCGYRCVNCAC 218

RESULT 2
ID 010614 PRELIMINARY; PRT: 151 AA.
AC 010614;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ACNPV ORF1 homolog.
OS Helicoverpa zea single nucleocapsid nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10468;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ELKAR;
RX MEDLINE=97321796; PubMed=9178498;
RA Le T.H., Wu T., Robertson A., Bulach D., Cowan P., Goodge K.,
RA Tribe D.;
RT "Genetically variable triplet repeats in a RING-finger ORF of
RL Helicoverpa species baculoviruses.";
RL Virus Res. 43:67-77(1997).
DR EMBL: U67264; AAB54095.1; -.
DR InterPro: IPR001841; Znf_ring.
DR SMART: SM00184; RING: 1.
SQ SEQUENCE 151 AA; 17548 MW; 4986432F6DCD3169 CRC64;

Query Match          69.1%; Score 56; DB 12; Length 151;
Best Local Similarity 70.0%; Pred. No. 0.086;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
   |||||:|:|
Db 101 CCGYKLCXHC 110

RESULT 3
ID 099H38 PRELIMINARY; PRT: 285 AA.
AC 099H38;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ie-0.
OS Helicoverpa armigera nucleopolyhedrovirus G4, and
OS Helicoverpa armigera nuclear polyhedrosis virus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=148363; 51313;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
RA Deng F., Chen X., Vlak J.M., Arif B.M., Hu Z.;
RT "Sequence analysis of the gp37 gene of Heliothis armigera single-
RT nucleocapsid nucleopolyhedrovirus.";
RL Zhongguo Bingduxue 15:35-42(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
RA Wang H., Hu Z., Sun X., Vlak J.M., Chen X.;

```

```

RT "Sequence analysis of the iap3 gene of Heliothis armigera single-
RT nucleocapsid nucleopolyhedrovirus.";
RL Zhongguo Bingduxue 15:43-49(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
RX MEDLINE=21078302; PubMed=11210934;
RA Wang H., Chen X., Wang H., Arif B.M., Vlak J.M., Hu Z.;
RT "Nucleotide sequence and transcriptional analysis of a putative basic
RT DNA-binding protein of Helicoverpa armigera polyhedrovirus.";
RL Virus Genes 22:113-120(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
RX MEDLINE=21064569; PubMed=11125177;
RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
RA Peters S., Zuidema D., Lankhorst R.K., Vlak J.M., Hu Z.;
RT "The sequence of the Helicoverpa armigera single-nucleocapsid
RT nucleopolyhedrovirus genome.";
RL J. Gen. Virol. 82:241-257(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;
RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
RA Peters S., Zuidema D., Lankhorst R.K., Vlak J.M., Hu Z.;
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=Helicoverpa armigera nuclear polyhedrosis virus; STRAIN=CL;
RA Zhang C.X., Wu J.C.;
RT "Genome structure and the p10 gene of the Helicoverpa armigera
RT nucleopolyhedrovirus.";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 33:179-184(2001).
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=Helicoverpa armigera nuclear polyhedrosis virus; STRAIN=CL;
RA Zhang C.X., Jin W.R.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF271059; AAG53751.1; -.
DR EMBL: AF303045; AAK96301.1; -.
DR InterPro: IPR001841; Znf_ring.
DR SMART: SM00184; RING: 1.
SQ SEQUENCE 285 AA; 33186 MW; C5FC3AE5BA27BDD CRC64;

Query Match          69.1%; Score 56; DB 12; Length 285;
Best Local Similarity 70.0%; Pred. No. 0.15;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
   |||||:|:|
Db 235 CCGYKLCXHC 244

RESULT 4
ID 08V5X4 PRELIMINARY; PRT: 285 AA.
AC 08V5X4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ORF8.
OS Helicoverpa zea single nucleocapsid nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10468;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen X., Zhang W.-J., Wong J., Chun G., Lu A., McCutchen B.F.,
RA Presnall J.K., Herrmann R., Dolan M., Tingey S., Hu Z.-H., Vlak J.M.;
RT "Genome sequence analysis of Helicoverpa zea single nucleocapsid
RT nucleopolyhedrovirus.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF334030; AAL56153.1; -.

```

DR InterPro: IPR001841; Znf_ring.
DR PROSITE: PS50089; ZF_RING.2; 1.
SO SEQUENCE 285 AA; 33189 MW; EF9E35A71B8E7F3D CRC64;

Query Match 69.18; Score 56; DB 12; Length 285;
Best Local Similarity 70.0%; Pred. No. 0.15;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
|||||:|
DB 235 CCGYKLCXHC 244

RESULT 5
O91B09 PRELIMINARY; PRT; 244 AA.

AC O91B09;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE ORF138.1e0.
OS Spodoptera exigua nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
NCBI_TaxID=10454;
[1]
PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=97437494; PubMed=9292027;
RA van Strien E.A., Faktor O., Hu Z.H., Zuidema D., Goldbach R.W.,
RA Vlak J.M.;
RT "Baculoviruses contain a gene for the large subunit of ribonucleotide
RT reductase.";
RL J. Gen. Virol. 78:2365-2377(1997).

RL J. Gen. Virol. 78:2365-2377(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20036646; PubMed=10567663;
RA Jukel W.F., van Strien E.A., Heldens J.G., Broer R., Zuidema D.,
RA Goldbach R.W., Vlak J.M.;
RT "Sequence and organization of the spodoptera exigua multicapsid
RT nucleopolyhedrovirus genome.";
RL J. Gen. Virol. 80:3285-3304(1999).

RL J. Gen. Virol. 80:3285-3304(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA Jukel W.F., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D.,
RA Goldbach R.W., Vlak J.M.;
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF169823; AAF33667.1; -;
DR InterPro: IPR001841; Znf_ring.
DR SMART: SM00184; RING.1.
SQ SEQUENCE 244 AA; 28719 MW; 1F7662E837A866DB CRC64;

Query Match 66.7%; Score 54; DB 12; Length 244;
Best Local Similarity 70.0%; Pred. No. 0.28;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
|||||:|
DB 199 CCGYKLCXHC 208

RESULT 6
O80L66 PRELIMINARY; PRT; 234 AA.

AC O80L66;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE TE0.
OS Mamestra configurata nucleopolyhedrovirus.

OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=191492;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=90/2;
RX MEDLINE=97163493; PubMed=9010313;
RA Li S., Erlandson M., Moody D., Gilloft C.;
RT "A physical map of the Mamestra configurata nucleopolyhedrovirus
RT genome and sequence analysis of the polyhedrin gene.";
RL J. Gen. Virol. 78:265-271(1997).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=90/2;
RX MEDLINE=21884635; PubMed=11886270;
RA Li Q., Donly C., Li L., Willis L.G., Theilmann D.A., Erlandson M.;
RT "Sequence and Organization of the Mamestra configurata
RT Nucleopolyhedrovirus Genome.";
RL Virology 294:106-121(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=90/2;
RA Li Q., Donly C., Li L., Willis L.G., Theilmann D.A., Erlandson M.A.;
RL EMBL: U59461; AAM09276.1; -;
DR SEQUENCE 234 AA; 27232 MW; 5E5F1330CD7A711 CRC64;

Query Match 63.0%; Score 51; DB 12; Length 234;
Best Local Similarity 60.0%; Pred. No. 0.81;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13
|||||:|
DB 193 CCGYKLCXHC 202

RESULT 7
O91BK7 PRELIMINARY; PRT; 289 AA.

AC O91BK7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Immediate early 0 protein.
OS Spodoptera litura nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G2;
RX MEDLINE=21425398; PubMed=11531416;
RA Pang Y., Yu J., Wang L., Hu X., Bao W., Li G., Chen C., Han H., Hu S.,
RA Yang H.;
RT "Sequence Analysis of the Spodoptera litura Multicapsid
RT Nucleopolyhedrovirus Genome.";
RL Virology 287:391-404(2001).

RL Virology 287:391-404(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=G2;
RA Yu J., Wang L., Hu X., Pang Y.;
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: AF325155; AAL01695.1; -;
DR InterPro: IPR001841; Znf_ring.
KW Zinc-finger.

SQ SEQUENCE 289 AA; 33387 MW; 50C3F3E63FE78C6B CRC64;

Query Match 60.5%; Score 49; DB 12; Length 289;
Best Local Similarity 50.0%; Pred. No. 2.1;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 4 CCGYKLCXHC 13

Db 245 CCGFRICNLC 254

RESULT 8

Q9BPE9 PRELIMINARY; PRT; 64 AA.
AC Q9BPE9;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Conotoxin scaffold IX.
OS Conus pennaceus (Feathered cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=37335;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conicello S.G., Gilad Y., Avdahan N., Ben-Asher E., Levy Z.,
Faltuzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides."
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF214980; AAG60408.1;
SQ SEQUENCE 64 AA: 6928 MW; 0AB87620FCCC1410 CRC64;
Query Match 59.3%; Score 48; DB 5; Length 64;
Best Local Similarity 60.0%; Pred. No. 0.81;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGYKLCXHC 13

Db 54 CCGYKLCXHC 63
ID Q950T3 PRELIMINARY; PRT; 514 AA.
AC Q950T3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 58.6 kDa protein.
GN C2861.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RC MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium."
RT Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Favello A.;
RT "The sequence of C. elegans cosmid C2861."
RT Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 RING-TYPE ZINC FINGERS.
DR EMBL; U41026; AAL02448.1;
DR InterPro; IPR000315; ZnF_Box.
DR InterPro; IPR001841; ZnF_Zing.
DR Pfam; PF00643; zf-B_box; 1.

DR Pfam; PF00097; zf-C3HC4; 2.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_2.
KW Hypothetical protein; zinc-finger.
SQ SEQUENCE 514 AA; 58600 MW; BC33388F0F599446 CRC64;

Query Match 59.3%; Score 48; DB 5; Length 514;
Best Local Similarity 63.6%; Pred. No. 4.9;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VCGYKLCXHC 13

Db 263 VCGYKLCXHC 273

RESULT 10

Q81680 PRELIMINARY; PRT; 112 AA.
ID Q81680
AC Q81680;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-NEO48;
RA TOKITA H., Okamoto H.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; D14198; BAA03233.1;
DR InterPro; IPR002166; HCV_RdRP.
DR Pfam; PF00998; HCV_RdRP; 1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
FT NON_TER 1
FT NON_TER 112
FT NON_TER 112
SQ SEQUENCE 112 AA; 12008 MW; D7D79CA732ED9D3F CRC64;
Query Match 58.0%; Score 47; DB 12; Length 112;
Best Local Similarity 66.7%; Pred. No. 1.9;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYCCGYKLC 10

Db 45 GYCCGYKLC 53

RESULT 11

Q39896 PRELIMINARY; PRT; 322 AA.
ID Q39896
AC Q39896;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE TGACG-motif-binding factor.
GN STF2.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC Eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-WILLIAMS; TISSUE-HYPOCOTYL;
RC MEDLINE=98388650; PubMed=9721678;
RX Cheong Y.H., Yoo C.M., Park J.M., Ryu G.R., Goeckjan V.H., Nagao R.T.,
RA Key J.L., Cho M.J., Hong J.C.;
RT "STF1 is a novel TGACG-binding factor with a zinc-finger motif and a
bZIP domain which heterodimerizes with GAF proteins."
RT Plant J. 15:199-209(1998).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY.

DR EMBL: L28004; AAC05018.1; -
DR TRANSFAC; T02973; -
DR InterPro; IPR004827; TF_bZIP.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
DR DNA-binding; Nuclear protein.
KW SEQUENCE 322 AA; 35238 MW; BE692E01EF6FA6E5 CRC64;

Query Match 58.0%; Score 47; DB 10; Length 322;
Best Local Similarity 66.7%; Pred. No. 4.8;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
DB 61 CGFPICHC 69

RESULT 12
039895

ID 039895; PRELIMINARY; PRT; 326 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE TGAGC-motif binding factor.

OS Glycine max (soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;

RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-WILLIAMS; TISSUE-HYPOCOTYL;
RX MEDLINE=98388650; PubMed=9721678;

RA Cheong Y.H., Yoo C.M., Park J.M., Ryu G.R., Goeckjan V.H., Nagao R.T.,
RA Key J.L., Cho M.J., Hong J.C.;

RT "STPL is a novel TGAGC-binding factor with a zinc-finger motif and a
RT bZIP domain which heterodimerizes with GBF proteins.";
RL Plant J. 15:199-209(1998).

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY.

CC EMBL: L28003; AAC05017.1; -.

DR TRANSFAC; T02969; -.

DR InterPro; IPR004827; TF_bZIP.
DR InterPro; IPR001841; Znf_ring.

DR Pfam; PF00170; bZIP; 1.

DR SMART; SM00338; BRLZ; 1.

DR SMART; SM00184; RING; 1.

DR PROSITE; PS00036; BZIP_BASIC; 1.

KW DNA-binding; Nuclear protein.
KW SEQUENCE 326 AA; 35488 MW; 0BF960434DC19AC5 CRC64;

Query Match 58.0%; Score 47; DB 10; Length 326;
Best Local Similarity 66.7%; Pred. No. 4.8;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
DB 61 CGFPICHC 69

RESULT 13
081547

ID 081547; PRELIMINARY; PRT; 364 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein (Fragment).

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-NE048;

RX MEDLINE=94201770; PubMed=8151307;

RA Tokita H., Shrestha S.M., Okamoto H., Sakamoto M., Horikita M.,
RA Iizuka H., Shrestha S., Miyakawa Y., Mayumi M.;

RT "Hepatitis C virus variants from Nepal with novel genotypes and their
RT classification into the third major group.";

RT J. Gen. Virol. 75:931-936(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-NE048;

RA Okamoto H.;

RL Submitted (MAY-1994) to the EMBL/Genbank/DBJ databases.

DR EMBL: D16613; BAA04035.1; -.

DR InterPro; IPR002166; HCV_RDRP.

DR Pfam; PF00998; HCV_RDRP; 1.

FT Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.

FT NON_TER 1
KW SEQUENCE 364 AA; 39979 MW; 9EE13125B350EF12 CRC64;

Query Match 58.0%; Score 47; DB 12; Length 364;
Best Local Similarity 66.7%; Pred. No. 5.3;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 GYCCGYKLC 10
DB 44 GLCCGYRRC 52

RESULT 14
09M1F9

ID 09M1F9; PRELIMINARY; PRT; 373 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical 42.7 kDa protein.

GN F9K21.60.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Jordan N., Bangert S., Wiedemann R., Voss H., Unsel M., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Queller F., Salanoubat M.;

RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL: AL138657; CAB75477.1; -.

DR InterPro; IPR002106; ATRNA_ligaset1.

DR InterPro; IPR002867; Znf_C6HC.

DR InterPro; IPR001841; Znf_ring.

DR Pfam; PF01485; IRR; 2.

DR SMART; SM00184; RING; 2.

DR PROSITE; PS00179; AA-trNA_LIGASE_II_1; UNKNOWN_1.

DR PROSITE; PS00518; ZF_RING_1; 1.

KW Hypothetical protein; Zinc-finger.

KW SEQUENCE 373 AA; 42688 MW; A4484B5BF0067BB CRC64;

Query Match 58.0%; Score 47; DB 10; Length 373;

Best Local Similarity 66.7%; Pred. No. 5.4;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 CGYKLCXHC 13
DB 61 CGFPICHC 69

Db 329 CGKFCYAC 337

RESULT 15

096BQ3 PRELIMINARY; PRT; 446 AA.
 ID 096BQ3
 AC 096BQ3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical 52.3 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL: BC015353; AAH15353.1; -
 DR InterPro: IPR001870; Gamma_carboxylase.
 DR InterPro: IPR003878; SPRY_domain.
 DR InterPro: IPR003877; SPRY_receptor.
 DR InterPro: IPR000315; znf_Box.
 DR InterPro: IPR001841; znf_ring.
 DR Pfam: PF00622; SPRY.1.
 DR Pfam: PF00643; zf-B_box.1.
 DR Pfam: PF00097; zf-C3HC4.1.
 DR PROSITE: PS00518; zf_RING_1; UNKNOWN_1.
 KW Hypothetical protein; zinc-finger.
 SQ SEQUENCE 446 AA; 52265 MW; 805E4AA25617724A CRC64;

Query Match 58.0%; Score 47; DB 4; Length 446;

Best Local Similarity 46.2%; Pred. No. 6.3; Mismatches 0; Gaps 0;

Matches 6; Conservative 2; Indels 0;

OY 1 VGVCCGYKLCXKC 13

DB 26 VVICGHSFCRPC 38

Search completed: December 2, 2002, 10:08:25
 Job time : 25.5 secs